DOCUMENT RESUME

ED 166 408 CE 019 498

AUTHOR	Wasik, J. L.
TITLE	The Use of Mathematical Models of Student Flow in Educational Planning: A Review and Critique. DASP Planning Paper No. 3.
INSTITUTION	North Carolina State Univ., Raleigh. Dynamic Analysis and Strategic Planning Program.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	75
CONTRACT	NE-C-00-3-0069
NOTE	58p.; For related documents see ED 133 479 and CE 019 499
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	Decision Making; *Educational Planning; *Enrollment Projections; *Mathematical Models; School Demography; Systems Approach

ABSTRACT

Designed for educational planners and decision makers, this report describes the application of mathematical procedures to depict the relationship of elements within models which represent the flow of students through an educational system. The introductory section includes a discussion of the history of the use of planning models in education and the steps in model development (variable identification, identification of variable interconnections, and operationalization and verification of the model). The next section reviews classification schemes for mathematical models of educational processes and presents a classification scheme developed for this report. Following this, examples of several classes of mathematical models of student flow are presented: extrapolation models; regression methods (simple first-order, polynomial, and multiple linear regression); and simultaneous equation models (non-optimizing models including the Markov Chain and optimizing models). Finally, the present utility of available mathematical models is discussed from the viewpoint of what effects these models have had on actual educational policy decision making. (Two related documents are also available in ERIC: Demography and Educational Planning [CE 019 499] and State and Local Responsibilities for Planning Occupational Education [ED 133 479].) (JH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED166408

THE USE OF MATHEMATICAL MODELS OF STUDENT FLOW IN EDUCATIONAL PLANNING: A REVIEW AND CRITIQUE

J.L. WASIK

****

The research reported herein was conducted pursuant to a contract with the National Institute of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily relecte official National Institute of Education position or policy.

****

DASP Planning Paper No. 3

Dynamic Analysis and Strategic Planning Program
Center for Occupational Education
North Carolina State University at Raleigh

1975

Contract No. NE-C-00-3-0069

CE 019 498

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John K. Coster

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

SUMMARY

Models of complex social systems can be represented concisely and unambiguously by the use of mathematical equations. For this reason, a wide variety of mathematical models have been developed to represent the flow of students through an educational system.

This report describes the application of mathematical procedures to depict the relationships of elements within models which represent the flow of students through an educational system. Examples of several classes of mathematical models of student flow are presented following a scheme developed for this review. A critical analysis is made of the features of an assumption underlying the reviewed models. Finally a discussion is presented of the present utility of available mathematical models from the viewpoint of what effects these models have had on actual educational policy decision-making.

CONTENTS

	Page
INTRODUCTION	1
A History of the Use of Planning Models in Education	3
Student Flow	5
Development of Models	5
Steps in Model Development	5
Variable Identification	6
Relationship Between Variables	6
Operationalization of the Model	6
Verification of the Model	6
CLASSIFICATION SCHEMES FOR MATHEMATICAL MODELS OF EDUCATIONAL PROCESSES	8
Mathematical Bases of Classification Schemes	8
A Proposed Classification Scheme	10
Further Observations Concerning the Classification of Mathematical Models of Student Flow	11
Level of Aggregation	11
Dynamic versus Static Models	12
Linear versus Nonlinear Models	12
MODELS OF STUDENT FLOW	14
Extrapolation Models	14
Model Basis	14
Variations in Ratio Type Models for Enrollment Forecasting	15
Regression Methods	16
Simple First-Order Regression	17
Polynomial Regression	19
Multiple Linear Regression	20
Simultaneous Equation Models	22
Non-Optimizing Simultaneous Equation Models	23
Markov Chain Models	27
Examples of Markov Chain Models	29
Optimizing Simultaneous Equation Models	34
APPROPRIATENESS OF MATHEMATICAL MODELS FOR STUDENT FLOW	37
Underlying Assumptions	37
Validation of the Model	38
Availability of Data	40

Page

Completeness of Model 41
Perceptions of Educational Planning Effectiveness 43

LIST OF REFERENCES 46

INTRODUCTION

Educational planning has been recognized as an important function of the local educational agency in the United States since the beginning of the 20th century. Moreover, the educational planning procedures used by educational systems or agencies have shown substantial increases in levels of sophistication during the past ten years.

Much of this interest in the use of educational planning procedures has been precipitated by the need for planning for changes in school operation mandated by legal and social pressures.[1] It has been noted that the age structure of the population in the United States is changing as the bulge in baby births has worked its way through the educattional system. As a result of changing public attitudes toward family size in the United States, birth rates are falling with the result that education can no longer be considered a "growth industry." Another factor which needs to be considered by the present-day educational manager is inflation, a phenomenon noted to be almost world-wide and likely to become a way of life. As a result of these outside influences which have traditionally been considered inconsequential to the operation of school systems, educational system managers are increasingly less certain about what the future will bring and what steps should be taken to plan for the future.[2]

Increasing attention is now being directed toward using sophisticated planning procedures to assist the educational system manager in the development of both long- and short-term educational programs. Implicit in the attempt to develop future plans is the need for a way to predict the future. It is obvious that performance of an educational system will be affected by the influence of elements outside the direct control of the system (e.g., local populational trends, inflation, legal directives) as well as those under the control of the school system (e.g., competence of instructors, adequacy of physical plant). Thus, any attempt to predict the future must take into account both of these types of influence. However, not much is known about the interaction of elements which influence the performance of the educational system either directly under the control of the educational system or not amenable to influence by the system.

---

[1]See, for example, the need for busing school children to meet school integration requirements or implementation of a school testing program to provide public evidence of the academic status and/or progress of students.

[2]While there is evidence that some of these trends were foreseen by researchers (see, for example, Cartter, 1971, on overproduction of scientific manpower), there is no evidence that such predictions were considered by educational system managers when making long-term plans.

Gass (1967) has advocated systems analysis as the approach to be used to gain a better understanding of the overall functioning of an educational system. Systems analysis has been described as a methodology which allows one to discover those properties which underly the operation of the system of interest.

Once the properties have been determined and the relationship between the properties noted, it becomes feasible to think about predicting what the outcome would be of some changes in the properties of the system of interest.

As Gass noted (1967, p. 143),

> . . ., educational planning can only mature on the basis of a much deeper understanding of the "internal dynamics" of the educational system, so that the implications of alternative policies can be explored, not only in terms of flows of pupils and students through the system, but also in terms of different structured alternatives, resource coefficients, and so on. In other words, there is no escape from the hard fact that policies for expanding the quantitative output of the system must be worked out on the basis of realistic assumptions about educational structure, alternative combinations of resources based on different technologies of teaching, changes in the curricula, and so on. An educational plan must in consequence incorporate policies towards educational innovation, and has to be looked upon as a tool of the overall policy for developing the educational systems.

Gass actually presented a use of model development which can assist educational researchers and planners to both understand the operation of an educational system and allow the prediction of performance of the system of interest. While total knowledge of the operation of an educational system is required for perfect prediction of system performance, it does not necessarily follow that increased knowledge of what comprises the system of interest will result in improved fidelity of forecasts. For example, knowledge of one element of a system will not result in increased predictive efficiency if the importance of this factor is due to its interaction with a second element which has not yet been isolated. Thus, while an educational system model may not represent the behavior of the educational system with total fidelity, it can still assist the educational planner in answering certain policy questions given that constraints on the use of such models are understood.

A History of the Use of Planning Models in Education

The Organization of Economic Cooperation and Development (OECD) formally proposed the use of systems analysis and operations research procedures in the allocation of educational resources in accord with national needs (Williams, 1969). Increasing or maintaining economic health is perceived to be an important function of the federal government, with the implicit assumption that it should be possible to maintain economic vitality by the use of planning procedures. This approach to educational planning was first presented at the 1961 OECD Washington Conference on Education and Economic Growth and is based upon the economic view that education should be considered "not only as a desirable consumption in its own right, but also as a necessary investment in future economic growth" (Gass, 1967, p. 140). Since education, unlike industry, does not operate in a free marketplace so that various alternatives for educational investments can be sorted out, economists feel the formulation of target goals for education investments is one of the primary functions of educational planning. A second and equally important aspect relates to a need for balanced resource allocation among the various levels of education (i.e., higher, secondary, and primary).

A corollary of this viewpoint is that growth can be insured if trained manpower will be available for the necessary expansion in production to maintain or improve the economic health of a nation. Thus, one of the requirements in this approach is an accurate forecast of the numbers of workers needed to maintain present levels of production or to man planned industrial expansion. Using manpower requirements, an educational planner working at the national or regional level derives estimates of the number of educational system graduates required to fill the new places in the industry. According to this approach, the educational planner then has the role of estimating the required numbers of educational systems graduates by field as a basis for rationally allocating resources throughout the educational system to meet the avowed educational goals. This approach has been labelled the manpower demand approach to educational planning. One reason for the acceptance of this approach in many OECD countries is that a federal government office is responsible for overall educational resource allocation, and, thus, funds can be allocated to institutions in accordance with an overall educational plan.

A second approach is based upon the concept of cost-benefit analysis or rate-of-return on investments in education. Educational planning, according to this view, not only considers the numbers of trained individuals required to meet national economic plan goals, but also decides whether or not some substitution of trained manpower could be made to achieve these same goals at a lower cost. While costs of providing trained manpower are relatively easy to measure, it is difficult to determine the benefits obtained from the use of manpower with alternative educational levels or even alternative educational background patterns.

The appeal of this view of educational planning lies in the concept that it should be possible to minimize the cost of producing the educated manpower required by a specific plan for economic growth. Unfortunately, the difficulty with estimating economic benefit has precluded the application of the cost-benefit concept in any situation other than as an academic exercise.

Educational planning has not, until relatively recently, been considered an appropriate responsibility of the U. S. government and, thus, has suffered from a lack of support at the national level. Papers by Kershaw and McKean (1959) and Fattu (1960) introduced the possible uses of systems analysis principles in the United States for the first time. Attempts to perform systems analysis of educational systems in the United States flourished briefly during the tenure of Dr. Alexander Mood as Assistant Commissioner of the U. S. Office of Education and Director of the National Center of Educational Statistics. In the middle 1960's, Mood established the Division of Operations Analysis. However, a later reorganization of the Office of Education resulted in the dismantling of this division and the stopping of all in-house work at the federal level.

The U. S. Office of Education has provided support indirectly for the use of operations research methodology in education by funding some educational research and development centers (e.g., the Western Interstate Commission for Higher Education). Several higher education institutions have supported, predominantly with state funding, the development of systems models to assist in institutional planning.

The role of the educational planner in the United States (and in Canada) at the federal level has been concerned with the development of educational enrollment forecasts, educational expenditures, and teacher requirements to maintain the educational system, rather than the development of plans to meet the manpower demands of the entire economic system. These actions are a function of the American--and other highly industrialized OECD member nations'--view that the individual's demand for education (i.e., social demand) is more important than state or national needs for trained manpower.

In summary, the major role of the educational planner is to provide the educational system manager with information on the estimates for demand for future educational services. Using either the manpower demand approach or the social demand approach to educational planning, an estimate of student numbers involved in various activities in the system is needed before rational decisions can be derived for resource allocation.

Student Flow

Accepting the premise of Gass that models of systems will most likely lead to an understanding of the operation of the system under study, it is necessary to recognize that complex systems such as an educational system at the national or regional (i.e., state) level can be represented as combinations of smaller systems which are referred to as subsystems. Thus, it is necessary to consider how the raw material is acted upon by a system until a final product is obtained. Students represent the raw material upon which the educational system is supposed to act. The final product will be the graduates of that particular educational system. Thus, to operate any model of an educational system, one must have a representation of the movement of students through the system. Once the structure of the system has been identified and the interrelationships defined, then one can study the flow of students through the system. One possible use of such student flow models is to provide predictions of student enrollment at the institutional level. A summary of mathematical models (including student flow models) for use in projecting enrollments in educational systems has been reported by Wasik (1971a). The present paper is specifically concerned with the representation of student flow and, as such, is particularly concerned with the development of a model which could stand alone or be used as part of a larger system.

Development of Models

Models can be represented at several levels of complexity beginning with the simple line drawing of the operation of a system and moving to a complex mathematical model which quantifies the relationships between the elements within the system of interest. While the use of a simple model such as a line drawing implies that the model developer understands the causes and effects which describe the operation of the system, rarely is it possible or even necessary to know the "educational physics" of the system to be studied (Correa, 1967a; Hammond, 1970) to develop a reasonable model. Thus, knowledge of the relationship between variables which give a good representation of system performance as opposed to perfect system behavior replication may be, in most situations, adequate. At this point it seems appropriate to describe the process by which models in general are developed. Later, discussion will be concerned with models used in educational planning and, more specifically, with models which develop student flow input for educational system models.

Steps in Model Development

Correa (1967a) has suggested that models have been often used in educational planning, but that the use of mathematical models in educational planning activities is relatively recent. He also posited that models are

"the building blocks of science." That is, by following the scientific methods utilized in the behavioral and social science research to study questions of interest, one is building a model. It is this reviewer's position that there is an essential difference between model-building in the scientific research sense and model-building in the planning sense. The former uses, in general, verbal statements to state the relationships of interest, while the latter uses mathematics to represent these relationships.

Variable Identification

The first step in model development is to identify the properties (i.e., variables) which will be used to characterize the state of each part of the process. Often the choice of variables used to represent a system has been dictated by the types of data that are available.

2. Relationship Between Variables

This step represents the identification of variable interconnections within the model. It is also at this step that attempts at noting the "educational physics" of the system of interest are made. Quite often at this stage, a sketch of the model is developed to assist the researcher in characterizing the flow of the model. This process allows one to categorize variables as dependent (i.e., a function of the other variables) or independent (i.e., a causal part of the model). It should be noted that independent variables can be exogenous (i.e., not influenced by the phenomenon studied) or endogenous (i.e., both cause and effect).

3. Operationalization of the Model

To operate the model, it is necessary to develop model mechanisms which will "mimic" the actual operation of the phenomena. Mathematics is a medium which allows the model to represent the actual system in operation. As will be noted later, a number of different mathematical techniques have been used to represent educational system models.

4. Verification of the Model

To perform a verification of any model, it is necessary to decide whether the model of interest should be accepted or rejected. To conduct the verification, the model developer will obtain information which is relevant to the system to be modeled. A typical approach followed in model verification is to obtain information on the variables included in the model at some past point in time for which data are available for both independent (model input) and dependent (model output) variables. The model then uses the independent variable data to predict the dependent variable values. The educational planner then decides whether obtained values from the model (i.e., the predicted values) provide a close enough fit to the already available dependent variable values. If the fit is close enough, the educational planner likely will decide the model can be made operational. A poor fit will send the educational

planner back through the model to see if some additional variable(s) should be included in the model or if some interrelationship(s) between variables should be changed. This process will continue until the planner has met an a priori stated criterion for model prediction or the project is terminated. A simple schematic showing the steps followed in model development for educational planning purposes is given as Figure 1.

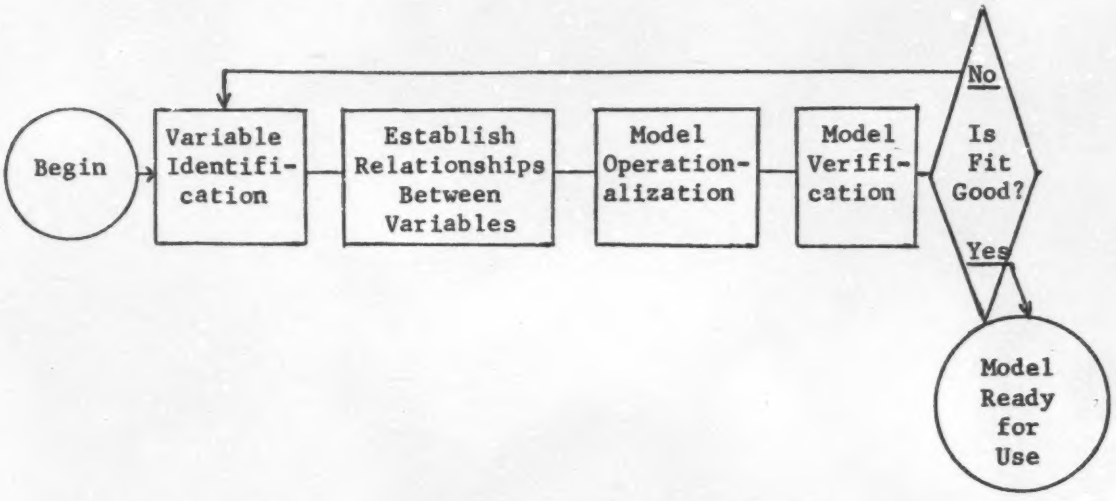

Figure 1. Steps Followed in Building a Mathematical Model

Model development and verification can be readily carried out if a model is comprised of quantifiable variables (i.e., variables which take a value). However, the process becomes substantially more complicated if any of the model variables has a qualitative form.

CLASSIFICATION SCHEMES FOR MATHEMATICAL MODELS OF EDUCATIONAL PROCESSES

A review of available education model classification schemes indicates that models are categorized with reference to two different dimensions: (1) model subject matter and (2) mathematical or computational structure. Schemes which have discriminated between models on the basis of their subject matter have been reported by Correa (1967a) and Wurtele (1967). Schemes which have concentrated on categorizing models according to the mathematical and/or computational structure have been presented by Hammond (1970) and Wasik (1971a,b). Johnstone (1974) joined these concepts for classifying mathematical models for educational planning in a scheme which categorizes models according to both model structure and model subject.

The availability of several different schemes for classification of mathematical models of educational processes suggests that there are different philosophical orientations implicit in each of the schemes. It is appropriate to consider these different orientations at this point.

Mathematical Bases of Classification Schemes

A common approach to classifying mathematical models is to consider whether the underlying model structure is deterministic or stochastic (Hammond, 1970).[3] This distinction was also suggested by other writers considering the classification of mathematical models of social science phenomena in general (Kemeny and Snell, 1962: Kendall, 1968; and Maki and Thompson, 1973). Hammond also recognized regression analysis as a technique which can be used to develop models but dismissed it on the grounds that the computational methodology allows for only an associational or relationship statement and not a cause-and-effect statement. This low opinion of the use of regression models for model development purposes is somewhat surprising in view of the wide use of regression analysis by econometricians in economic model-building. See Ball, 1968, for an exposition of the use of regression analysis in model-building.)

It might also be noted that Hammond considered this procedure to be a statistical technique and not purely mathematical and, thus, not appropriate for use in classifying mathematical models of educational systems! He did, however, indicate that other model attributes can be used to characterize the model. Among the characteristics he considered

---

[3]Another way of attempting to denote this difference in model bases is exemplified by the use of simple linear regression equations. If one is interested in developing an estimate of enrollment in a specific university course as a function of total university enrollment, one could

important for describing a model are scope and complexity of the model, model inputs and outputs, and whether a model is descriptive or normative.

Wasik (1971a, b) used the term model structure for distinguishing between mathematical models of educational systems. Models classified according to this scheme are distinguished on the basis of their complexity (and scope), an attribute presented by Hammond as appropropriate for use in model description. Wasik's categorization scheme provided for three levels of classification of models: (a) extrapolation, (b) structural, and (c) Markov Chain-type. Models in this system are classified on the basis of whether one equation was used to generate one output measure, i.e., extrapolation; several equations were used to generate one or more outputs, i.e., structural; or matrix multiplication was used to generate several outputs, i.e., Markov Chain-type.

Johnstone (1974) categorized mathematical models into four different mathematical approaches: deterministic, Markov Chain, regression and mathematical programming. An inspection of the types of models included by Johnstone under the deterministic category shows a high degree of similarity to models classified by Wasik as extrapolation and structural models. Actually, Johnstone might have been on firmer ground if he had stated that his classification dimension represented compuational approaches. For as he recognized, once the mathematical assumptions including the Markov Chain theory are relaxed, the same computational results will be obtained with a deterministic difference equation approach. Also, the mathematical programming approaches cited by Johnstone as presently being used in educational planning have a deterministic mathematical basis and, thus, could logically be grouped as deterministic models.

Other attributes suggested for use in categorizing mathematical models are whether the model is static or dynamic and whether the variables in the model are linear or nonlinear (Beach, 1957).

While Wurtele (1967) specifically stated he would not use mathematical form as a way to classify models, he did cluster the mathematical methods used to represent educational systems or educational system components in terms of the underlying computational approach, e.g., demographic models based upon transition matrices in either macro or micro form, trend extrapolation, cost-benefit (described as mathematical programming models by Johnstone), and simulation methods. In contrast, Correa's structure, which is the most inclusive since it allows for

---

develop a regression equation relating past enrollment in the course to corresponding school enrollment. Then, if one uses an estimate of the next year's total institutional enrollment to give a single "point" measure of expected course enrollment, the obtained regression equation will be used as a deterministic model. In contrast, if a confidence interval is developed about the predicted value obtained from the regression equation, then an individual is 95 percent confident the actual course enrollment will fall within the calculated confidence interval, and the model is being used in a stochastic fashion.

levels of model attribute categorization, is solely concerned with considering the purpose for which the model was developed. An analysis of the scheme does show that Correa distinguished betweeen models which optimize between alternatives (i.e., mathematical programming) and models which do not optimize between alternatives.

An analysis of the various classification schemes described above does not allow one to identify a widely accepted scheme for classifying mathematical models of educational systems. Available classification systems have used both mathematical assumptions and computational procedures as a basis for categorizing models. Moreover, it appears that the present use of mathematical procedures to model educational systems is a function of the modeler's familiarity with a particular mathematical and/or computational approach and not the result of the logical development of an equation or system of equations to represent the operation of the system of interest. It appears that this situation arises more from the failure to use the research methodology of applied mathematics than from a lack of understanding of mathematics.

A Proposed Classification Scheme

A logical approach to devising a particular classification system of mathematical models for educational planning purposes preserves the distinction of the deterministic versus stochastic bases of mathematical models within the classification scheme. The use of computational approches as a means of classification is implicit in the Johnstone approach (1974). However, a major difficulty in the development of a classification scheme arises from the fact that educational planning models can utilize different mathematical approaches within the same overall operational model. Kendall (1968, p. 2) referred to these models as "determino-stochastic" models since they include elements of both the deterministic and stochastic types of models. This situation can arise when a model of a complex system is developed in a piecemeal fashion such that each subsystem of the total social (i.e., educational) system is represented by a different mathematical model. In a case of this type, it is quite possible that the different subsystem models may be developed from different mathematical assumptions which, thus, could lead to such mixed models. For example, Wasik (1971b) developed a model for predicting enrollments in a state community college system which required a regression model to forecast total curriculum enrollment and a Markov Chain-type model to use the total enrollment forecasts and distribute expected numbers of students among different curriculum areas. It seems essential that any classification scheme used for classification of mathematical models provide for unambiguous classification of elements when all important model attributes are considered.

Considering the issues raised in the above review, a strong case can be made for a two-level classification system in which the major dimension partitions models according to model structure, and specific model forms represent a nested subclassification. This classification follows Wasik's (1971a, b) formulation which considers models to be extrapolation (i.e., ratios or regression) and simultaneous equation (i.e., structural). Extrapolation models are subclassified by whether they represent ratio or regression-type models. Simultaneous equation models are further subdivided into models which allow for optimal solutions subject to constraints (i.e., mathematical programming models) or nonoptimizing models (i.e., Markov Chain-type).

Further Observations Concerning the Classification Of Mathematical Models of Student Flow

A basis for categorizing mathematical models of educational systems has been described above; however, there are other characteristics of models which have been suggested as being relevant to the description of mathematical models in varying degrees (Beach, 1959; Hammond, 1972). The following sections will consider some of these aspects of mathematical models not specifically covered by the classification system presented in Figure 1.

Level of Aggregation

According to Correa (1967a) and Wurtele (1967), educational models which consider individual student characteristics in the operation of the model are micro models; macro models represent elements as a specific aggregated level (i.e., classroom of students, total grades of students, schools of classrooms, etc.) While Correa specifically limited micro models to those of the psychological processes representing learning, Wurtele's classification scheme allows for consideration of models in which the individual subject is the basic unit model of an educational system.

It seems reasonable to expect that further development of educational models should include a model of the individual student's learning process as a subsystem within a model of an educational system. One example of this approach can be found in McReynolds (1971). He presented alternative models which have utilized psychological prediction procedures for indivdual students as a subsystem within a model designed to estimate future enrollment within the Ontario, Canada, post-secondary system. While he presented data which showed a fairly good fit from the predictions to the available data, there has been no further development of this innovative and promising approach.

Dynamic versus Static Models

This review is concerned with mathematical models of educational systems which provide either directly, as output, or indirectly, as as intermediate model data requirements, a representation of students moving through an educational system. Models which focus particularly on student movement through a system are generally labeled as student flow models. The terms "flow" and "movement through a system" suggest a dynamic system. While it is fairly easy to envisage how simultaneous equation models can be developed to represent a dynamic situation, it may be more difficult to see how extrapolation models could be used to represent elements within a dyanmic system.

It seems to this reviewer that the use of successive iterations of the extrapolation equation for different activities at several points in time can provide a reasonable representation of how students may flow through a system. This simple adaptation of a single output static model actually mirrors the process followed in the mathematical calculations resulting from the use of a Markov-type model. By the enumeration of obtained values of expected number of students for each educational state, one could logically argue that student flow was occurring. The most obvious deficiency in such an approach to model-building is that joint effects resulting from the interconnection of factors not immediately connected by the model in terms of time and/or location may result in substantial errors in enrollment estimates. That is, the advantage accruing from the use of dynamic models is that the various states of the model are responsive to the joint effect of factors occurring at some point in time removed from the particular educational state for which it is desired to obtain an estimate of enrollment in much the same way the use of feedback loops in a cybernetic system allows a system to modify a prediction upon receipt of new information.

Linear versus Nonlinear Models

The form of the mathematical equations used to represent the educational system of interest becomes the actual mathematical model. Specification of variables and/or parameter values is also a part of the development of a mathematical model. In general, most mathematical models assume that linear relationships exist between variables included in the model, or at least that the relationships can be transformed mathematically so that fairly straightforward algebraic or numerical procedures can be used to generate solutions. However, a number of researchers concerned with the development of models of real world systems have doubts that linear models of the real world system will suffice; it is their contention that the real world is nonlinear (Sisson, 1968; Lucas, 1974). The real difficulty in developing nonlinear models is computational; procedures available today are not good enough to obtain answers because of the complexity of the systems of equations. At this point it seems that there are two alternative paths which can be followed: (1) The researcher can begin with a simple model which is linear in the parameters and attempt to develop some degree of predictive efficiency before adding nonlinear components in an attempt to improve

the model's prediction. Inclusion of a few nonlinear components in the model at one time should considerably improve the chances that the model can be solved by presently available analytical techniques. (2) In contrast, a second approach to getting around the computational problem is to use computer simulation procedures to represent the operation of the educational system. Thus, the problem of mathematical intractability for solving the system of equations is avoided.

The following sections will review, in turn, mathematically based models for representing student flow through educational systems by considering extrapolation models and simultaneous equation models. Subcategories of models under each model structure classification will also be provided. In each case care will be taken to indicate whether the proposed model has an underlying deterministic and/or stochastic basis. Included within this review will be models which have been proposed and not implemented as well as those student flow models which have been in operation for purposes of assisting educational planners in studying the operation of educational systems. In addition, methods which have been noted to be useful in contexts other than educational planning will be considered. Finally, a summary of advantages and disadvantages of using various mathematical models described in the review will also be presented.

MODELS OF STUDENT FLOW

The structure of this review of models useful in representing student flow through an educational system is as follows. Examples of the various modeling approaches will be presented separately under the appropriate major category of model structure (i.e., extrapolation or simultaneous equation system). At this point, a description of the mathematical form and operation of the models will be given. Descriptions of models will also indicate whether the model appears to have a deterministic or stochastic mathematical basis and whether the model is at a macro or micro level of aggregation of system elements. Finally, any other relevant information which may also distinguish a particular model from others in its class will be presented.

Extrapolation Models

The basic approach to developing extrapolation models is presented first, with variants of the basic model presented in later sections.

Model Basis

Extrapolation models as used in educational planning are almost entirely deterministic in nature. The development of enrollment forecasts of a school grade (class) for a particular year can be used to demonstrate the essential nature of extrapolation models. Assume first of all that information regarding the particular school and its associated attendance zone (area from which the school draws its students) could include data on such variables as:

1. previous grade's enrollment;
2. present grade's enrollment;
3. proportion of students who pass to grade i from grade i-1;
4. proportion of present year's grade enrollment held back;
5. proportion of school enrollment likely to "migrate out" (i.e., move out of school district);
6. number of students likely to "migrate in" to school district; and
7. total population served by school.

One simple model which is commonly used in enrollment forecasting is the cohort-ratio (cohort-survival) model. A mathematical representation of this model is given as (1).

$$(1)\quad y_t = p_i\, y_{i,t-1}$$

where $y_{i,t}$ = enrollment in grade i for year t

and p_i = proportion of students passing to grade i from grade i-1.

Adair (1969) presented a description of how to implement such a model at the local school system level. He suggested calculating a survival rate for grade i by calculating the ratio of number of students in the ith grade at time t to the number of students in grade i-1 at time t-1. These survival ratios for a number of years are then averaged to provide mean survival ratios. This procedure can be carried out for each of the grade levels in the school system by the use of the relationships given in equation (1) so that a forecast of enrollments for grade i at time t is found by multiplying the number of students in grade i-1 at time t-1 by the average ratio for grade i.

A more sophisticated ratio-cohort model makes use of information on the number of students who will migrate into or out of the particular school district. This model is presented as (2):

$$(2)\quad y_{i,t} = p_i y_{i,t-1} + q_i y + 1.$$

In this case, the equation represented as (1) is supplemented by the number of students in grade i obtained from the previous year as calculated by the product of the retention rate for grade q_i (i.e., $q_i = 1 - p_{i+1}$) and the enrollment in grade i for the previous year (i.e., $y_{i-1,t-1}$). The term I represents the difference between in-migrating and out-migrating students for that system appropriately aged for grade i. Migration figures can also be based upon past data available for migration rates and total population in the school district.

Variations in Ratio Type Models for Enrollment Forecasting

While the relatively straightforward approach implicit in the survival-cohort method has found extensive use throughout the world by dissemination through the offices of UNESCO (Jacoby, 1959; Liu, 1966) and OECD (1967), there have been continuing attempts to improve the efficiency of prediction resulting from the use of such a model. Lins (1952) has presented the use of the Class-Rate Progression Technique as a variation in the cohort-ratio procedure. In this case, specific individuals are followed from one educational state to another in the educational progression sequence so that the obtained ratio of class promotion is based upon single samples. While the method is mathematically more complex than the cohort-ratio procedure, surprisingly accurate results have been obtained by the use of this procedure at the University of California at Los Angeles for course enrollment forecasts (Smith, 1970).

Other variants of the cohort-survival procedure have been based upon the inclusion of information on potential students as estimated by information on the numbers of pre-school (i.e., ages 1-4) or out-of-school youth obtained either by survey or by use of census information. For example, Cross and Sederberg (1969) utilized a computerized model to give Minnesota school district administrators estimates of future school enrollment in their school district by grade. While their model was generally based upon cohort-survival rate information, they also used

an estimate of numbers of students entering school for the first time to provide for first-grade enrollment forecasts. These potential first-time enrollment estimates were determined by reference to census tract enumerations of numbers of children residing in a particular school district.

McNamara (1971c) suggested the use of a modification of the cohort-survival method obtaining enrollment estimates for each of the grades in a school. By placing a larger weight on the more recent information of enrollments per grade, he felt the process of forecasting would be improved. No data were presented, however, to provide support for his claim of increased forecasting efficiency by use of this wieghted cohort-survival procedure. All of the above models can be described as deterministic with regard to assumptions and operating at a macro level.

An evaluation of the effectiveness of the cohort-survival method for use in predicting school district attendance for 242 communities in the State of Massachusetts was reported by Grennawalt and Mitchell (1966). Using information available from the Massachusetts State Department of Education for the years 1952 to 1959, accuracy of predictions obtained by successive applications of the cohort-survival method was compared for each of the 242 communities for the period of interest. Community growth factors were also related to predicted and actual enrollment figures. It was found that roughly 60 percent of the forecasts exceeded a 10 percent error in forecasting. Forecasting was found to be relatively better for larger communities than for smaller ones and was likely to underestimate the actual school enrollment in large community schools and to overestimate the actual enrollment in small community schools. (An exception to this rule was found for the 5,000-15,000 population size communities, where underestimation was also noted.) Further, it was noted that as changes in community population and school population occurred, larger discrepancies between predictions and actual enrollments occurred. A final observation was made that even communities which had accurate (i.e., 10 percent or less discrepancy) predictions still might have large under- or over-prediction if the population base of the community was large enough.

Regression Methods

The computational procedure utilized to obtain enrollment forecasts by regression is similar in operation to the cohort-survival ratio methodology and its variants. The major difference lies in how the parameter values used in the prediction equation are derived. As noted earlier in the discussion of the mathematical bases of educational planning models, the regression model is based upon statistical theory which posits a stochastic process as representing the linear model. However, once the form of the equation has been derived through the point estimate of coefficients associated with each of the variables,

simple arithmetic computations only are required to estimate enrollment much as in a simple extrapolation procedure.

Following a general format for writing regression equations developed by Draper and Smith (1966, pp. 128-129), the regression analysis can be written as a linear model in the independent variable X_1, X_2 . . . X_k in the form:

$$(3)\quad Y_i = \beta_o Z_{io} + \beta_1 Z_{i1} + \beta_2 Z_{i2} + \ldots \beta_p Z_{ip} + \varepsilon_i$$

where β_o = the true value of the intercept,

Z_{io} = a dummy variable taking the value 1 for every element,

β_j = the true slope of the relationship between Y and X_j,

Z_{ij} = a general function of X_1, X_2 . . . X_k; that is,

$Z_j = f_j (X_{1j}, X_{2j} \ldots X_{kj})$ and Z_{ij} can take any form including involving only one X variable, and

ε_i = random error, normally and independently distributed with a mean of 0 and a variance of σ^2.

The use of regression analysis in educational planning enrollment prediction has resulted in three general types of approaches: (1) simple linear (first-order), (2) polynomial (second-order or above), and (3) multiple linear regression (k independent variables which can have either linear or polynomial forms). The mathematical form of the linear model and a discussion of the application of the method to enrollment forecasting in educational planning will be presented for each of the three cases.

Simple First-Order Regression

The general form for simple first-order models with one independent variable is

$$(4)\quad Y_i = \beta_o Z_{io} + \beta_1 Z_{i1} + \varepsilon_i$$

where the terms are defined as in equation (3).

One widely used regression analogue to the cohort-survival procedure, equation (1), is the cohort-regression method. The equation in the cohort-regression model, written in simple form, is presented in equation (5).

$$(5)\quad Y_{i,t} = C_{o_{it}} + P_{it}\, Y_{i-j,t-j} + \varepsilon i,t$$

where $Y_{i,t}$ = enrollment in grade (class) i at time t,

$C_{o_{it}}$ = a constant intercept term,

P_{it} = the sample estimation of the true relationship between enrollment in year t and year t-j, and

$\varepsilon_{i,t}$ = random error associated with the i^{th} observation for the t^{th} year

The idea is to predict enrollment in year t with knowledge of numbers of potential students j years previously. Two variants of this approach were utilized in the methodological study on prediction of educational attendance conducted by the Department of Statistics of the University of Minnesota (Brown and Savage, 1960). A live-birth method was developed to predict the number of high school graduates from the Minnesota public schools, given the number of live births 18 years previously. In equation (4), the index j would be set equal to 18 to indicate the average number of years which elapsed between birth and graduation from high school. The second model used a regression equation to predict numbers of students in a particular grade by utilizing $P_{i,t}$ as an index of the tendency of students to pass from grade i-1 to grade i. Thus, j = 1 in the case of predicting grade-to-grade promotion. The number of live births would be used for predicting first-grade enrollment, as is done in the first Minnesota model, and j = 5.

To use these approaches to enrollment forecasting, the number of births observed in year t-18 is multiplied either by the regression coefficient associated with the ratio of graduates at time t to births at time t-18 (Brown and Savage, 1960, Model I) or by the product of transition rates of promotion from one grade to another through graduation (Brown and Savage, 1960, Model II) to obtain the expected number of graduates from Minnesota public high schools in year t. A comparison was made of estimates obtained under Models I and II with estimates derived independently by the Minnesota State Department of Education. An analysis of these estimates revealed discrepancies, with the State Department estimates observed to be persistently higher. Since actual data were not available at the time of publication of the report, it was not possible to determine the accuracy of any of the models.

Webster (1970) reported the results of a study which determined the validity of enrollment projections obtained under the cohort-survival ratio method and the regression approach devised by Brown and Savage. Utilizing data available for a five-year period for 25 Michigan school districts stratified by the factor of student enrollment growth rate, 13 survival ratios for each of the transitional grade-to-grade

progressions were calculated and then utilized to predict enrollment for elementary and secondary grades for five additional years.

The linear regression approach simply developed separate prediction equations for each grade level. The predicted enrollments for elementary grades (1-8) and secondary grades (9-12) were then summed to give separate elementary and secondary totals for each method. A ratio of the difference between predicted and observed enrollments divided by observed enrollment was calculated as an index of fit. Regression analysis provided a better estimate of enrollment totals in 18 of the 25 cases, an outcome statistically significant at the .15 level. It should be noted that Webster's regression procedure paralleled the cohort-survival approach in that regression equations were calculated separately within grades.

Other first-order regression approaches have been based upon the use of curve fitting techniques. These procedures are based on growth-type curves which require linear transformation of the equations to allow use of linear regression methodology. The three most commonly presented for possible use are (1) Exponential Curves, (2) Gompertz Curves, and (3) the Logistic Curve. Kuang (1955) presents a discussion of the use of curve fitting in estimating enrollment, but he does not provide data on which a basis for evaluating the effectiveness of these methodologies for forecasting.

Polynomial Regression

Polynomial regression can be considered simply an expansion of the simple linear regression model to include base enrollment values (i.e., $Y_{i-j,t-j}$) raised to the second order or higher. A general form of the equation is given as equation (6).

$$(6)\quad Y_{i,t} = Co_{i,t} = P1_{i,t}\, Y_{i-j,t-j} = P2_{i,t}\, Y^2_{i-j,t-j} + \cdots P_{k_{i,t}}\, Y^k_{i-j,t-j} + \varepsilon_{i,t}$$

where $Y_{i,t}$, $Co_{i,t}$ and $\varepsilon_{i,t}$ are defined as in equation (5), and

$P_{1_{i,t}} \cdots P_{k_{1,t}}$; i = 1, 2, . . . k represents the sample partial regression coefficeint of the enrollment in year t-j for grade i raised to the i^{th} power after controlling for effects of enrollments raised to powers other than i.

Applications of the use of polynomia regression have been presented by Sawhiris (1970) and Impara (undated). Sawhiris was interested in the estimation of individual course enrollments for an educational psychology course at Sacramento State College in California. His model III used a polynomial form in an effort to capture the trend of enrollments over the eight previous semesters of enrollment in that course. Using the average residual squared (i.e., S^2) as a criterion of effectiveness, he concluded that it was as effective as other types of multiple regression models (to be discussed in the next section) for forecasting enrollment.

Impara was interested in the development of a computer-assisted long-range planning model for forecasting educational needs in Florida. In the development of enrollment forecasts for grades 1-12, he found that cohort-survival methods provided the best prediction, except at the first-grade level where a polynomial model to the second order using births six years earlier was the best.

Multiple Linear Regression

These models are similar to those in the first and second classes from the viewpoint that the model can include both first-order and higher-order effects for more than one independent variable. This general class of models also allows the includsion of cross-product terms where the generated independent variable is a product of two or more first- or higher-order independent variables. The majority of models classified under this subheading can be referred to as econometric models. Thus, a multiple linear regression model can include lagged variables. (i.e., the dependent variable measured at some earlier point in time) as well as various cognitive, economic and social variables which may be hypothesized to have an effect on whether or not a student (or group of students) will enroll in a particular grade. Equation (7) presents a general econometric model.

$$(7)\ Y_{i,t} = \beta_{o_{i,t}} + \beta_{1_{i,t}} Z_{i-j,t-j} + \beta_{2_{1,t}} Z_{1_{i,t}} + \beta_{3_{i,t}} Z_{2_{i,t}} + \ldots \beta_{k_{i,t}} Z_{k-1_{i,t}} + \varepsilon_{i,t}$$

where the terms are defined as in those models represented in equation (3).

The use of multiple linear regression procedures to predict enrollment was also used by Sawhiris (197) to provide a contrast to the polynomial model reported earlier. In his attempt to forecast enrollment in the educational psychology course at Sacramento State, he developed two addtional models. Model I included linear and quadratic time effects and a semester (i.e., fall or spring semester) effect to predict attendance in the course. Model II included an effect for overall trend, semester, and previous semester's enrollment as a

lagged variable. A comparison was made of the predictive efficiency of these two models with the polynomial regression model (Model III), using an average of the residual squared (S^2) as a criterion of effectiveness. Models II and III were found to be equally effective and somewhat more effective than Model I in predictive efficiency.

One example of an iterative process for model development using regression analysis is provided by the second phase of the University of Minnesota educational attendance study. One of the stated purposes of the study was to provide university enrollment forecasts for use in educational planning and institutional development. Two models were posited as beginning points in the process to develop an enrollment prediction model. The first model used a time series function to fit available data for the years 1921 to 1960. A second model simply predicted university enrollment as a function of numbers of Minnesota high school graduates. It is of interest to note that this model would use as input the predictions obtained from simple cohort-regression models discussed in the previous section. Thus, we have an example of two subsystems (i.e., predict high school graduates - predict university enrollment) being used to develop a single, overall model for university enrollment.

The researchers working on this phase of the project concluded that the two original models were about equally effective in predicting enrollment, but that some estimates provided substantial deviation from actual values. Since this lack of fit was associated with the war years, a variable was added into the second model to take into account the national military requirement changes. Evaluation of this model showed a smaller average squared difference (S^2) but a fit that was markedly poorer over a wider period of years. Another model was developed to include lag in student enrollment at the university to account for students who may take a year off from their studies before enrolling at the university. Thus, the fourth model included high school graduates for year t and year t-1 (the lagged variable) and military personnel requirements for year t. The resultant mean square error was less than one-half the mean square error obtained for the previous model, thus demonstrating the success of this approach. One other model was developed solely for the purpose of projecting freshman and upperclassman enrollment by separate equations. By aggregating the numbers of freshmen and upperclassmen, it was possible to obtain total institutional enrollments as was done in Model IV directly. While this procedure would appear to be most likely to provide the best estimate (due to the fact that separate equations for freshmen and other students should account for possible disturbance), a larger observed S^2 for total enrollment indicated this approach to be less accurate.

A similar iterative model development process was followed by Haggstrom (1969) in his attempt to obtain total estimates of demand for future higher education and the costs of meeting these demands for the United States. Utilizing a logistic curve, he derived separate models which seemed to fit well for both male and female enrollments in the years prior to 1965 except for the males where large numbers of prospective students (i.e., men aged 18-21) were in the armed forces and/or had recently returned to postsecondary study. By including variables such as high school class size, number of draftees, and number of veterans

studying under the GI bill, a much better fit of the predicted to actual enrollment was obtained. Other multiple regression equations have been developed which utilize economic variables in addition to such obvious variables as high school graduates (to predict university attendance) or total population (to predict grade and high school attendance). Banks and Hohenstein (1970) made comparisons among the cohort-survival, cohort-regression, and multiple regression approaches for predicting national higher education total enrollment. They found the multiple regression model, which included the following exogenous variables--(1) deflated family income, (2) consumer price index, and (3) total high school graduates--to be markedly superior in prediction (i.e., $R^2 = .99$) to the other methods for estimating enrollment in the 1960's. In the original model, total first-time freshman enrollment and current year's enrollment were included as variables but were found to be nonsignificant predictors. While they suggested there were difficulties in attempting to estimate the deflated family income and consumer price index for future years, Banks and Hohenstein suggested some methodologies for estimating future falues of these variables.

Another example of this type of model was presented by Wasik (1971a), who was interested in projecting enrollments of community colleges. His model included, as exogenous variables, regional (i.e., county) population, military needs on a present and lagged basis, and an index of economic activity. His rationale for including economic activity as a variable was based upon the observation that enrollment in a postsecondary institution seemed to be negatively related to the amount of economic activity in a district; that is, attending a community college was a reasonable alternative for individuals who could not get a job.

Simultaneous Equation Models

The above treatment of models has been concerned with single equation methods for developing enrollment projections. While it is possible that one equation may give information of use in obtaining student projections, it is more likely that the enrollment process may be only one part of an interconnected system of relations which describe the operation of an educational system. Thus, more than one equation would be required to depict the system in operation. Models which include sets of equations to describe model structure are referred to as simultaneous equation models. In the discussion to follow, the mathematical bases of simultaneous equation estimation procedures for representing student flow are presented first, followed by short descriptions of various types of simultaneous equation models.

Two major classes of simultaneous equation models are identifiable. Following Correa's notation, they are models which do not provide a choice among alternative policies (non-optimizing models) and models which require than an optimum choice among alternative policies (optimizing

models)[4]. Non-optimizing models will be treated first in this section of the review.

Non-Optimizing Simultaneious Equation Models

Simultaneous equation models have been developed for use in the development of enrollment forecasts by class (grade) or student characteristic (graduated versus still-in-school versus not-in-school). The approach is to develop a set of equations which represent the expected numbers of students in the system being modeled. The Markov Chain model is the mathematical model most widely used for educational planning in general and student flow in particular and, thus, will be considered separately as a specific form of non-optimizing simultaneous equation models.

Non-optimizing simultaneous equation models are generally deterministic in nature and are comprised of sets of difference equations. Characteristics of these difference equation models include specification of linearity in the parameters, first-order difference operations, and constraints in model coefficients. That is, the variable values used to generate enrollment forecasts are set in advance of running the model. Either constant values or a specified function of variable values can be used as input into the model. This type of approach also assumes no changes in the value of the input variable coefficients will occur in the future.

A straightforward example of the use of simultaneous equation models can be found in the work of Bolt, Koltun, and Levine (1965) on modeling doctoral scientist and engineer feedback into higher education. They utilized the following two linear first-order difference equations to represent the behavior of doctoral graduates in seeking positions in year k:

$$(8)\quad E_{k+1} = (1 + b_r - a - c)\, E_k$$

$$(9)\quad P_{k+1} = (1 - a)\, P_k + [(1 - b)\, r + c]\, E_k$$

where E = number of doctorates who serve relevant educational functions (i.e., teach or do research),

P = number of doctorates who are engaged in professional activities other than education,

[4]Non-optimizing models allow for study of future evolvement of educational systems under a number of different policy alternatives. The educational planner can then use these results to decide which policy will meet stated goals of the educational system. An optimizing model will give a solution that will directly indicate what policy should be adopted to attain the desired educational system goals.

a = attrition rate for doctorates (i.e., leaving profession),

b = feedback ratio for those doctorates who return to education in either teaching or research,

c = transfer rate of doctorates who transfer from educational to non-educational activities, and

r = production ratio (number of doctorates awarded that year divided by number of doctorates performing educational activities).

This set of two simultaneous difference equations has the unique solution:

$$(10)\quad E_k = (1 + b_r - a - c)^k E_o$$

$$(11)\quad P_k = (1 - a)^k P_o + \left[(1 - b_r - a - c)^k - (1 - a)K \right] \frac{[(1 - b)\, r + c]}{br - c} E_o$$

The sum of quantities calculated by equations (10) and (11) gives the total number of doctoral scientists and engineers for year k.

Simultaneous equation methods have found wide range in economic planning. Utilizing the characteristics of a system from an economic viewpoint, it is possible to calculate the production of doctorates by equation (12),

$$(12)\quad \alpha = r\left[b - \frac{a + c}{r}\right],$$

the damping factor of production by equation (13),

$$(13)\quad \beta = \frac{1 - \alpha,}{1 + \alpha}$$

and the equilibrium point between doctorates in the other-than-education sector and doctorates in the education sector by equation (14),

$$(14)\quad \gamma = \frac{1}{b - c/r} - 1.$$

This model is particularly noteworthy in that it provides for specific educational mechanism which assumes that demand (applicants) for higher education is greater than supply (spaces in universities).[5]

[5]This assumption is still realistic today; however, it is likely that the reasons for this situation have changed. Graduate departments at many universities have begun to admit smaller numbers of students to their programs on the premise that the job outlook is not as bright as a few years ago.

A more extensive model for estimating educational enrollment was developed by Tinbergen and Bos (1965a) for the explicit purpose of attempting to coordinate the educational system's production with the needs of economic development for a particular country. This model was designed to project the flows of educational manpower which would be available to an expanding economy given an expansion also occurred in the educational system. This model contained six equations which represented the relationships among production, the labor force, and the educational system; these equations were solved simultaneously to give the expected educational output necessary to attain the specified levels of economic activity.

The Tinbergen-Bos approach was adopted by the OECD for use in the Mediterranean Project. Countries involved were Spain (Emmerij, 1965), Turkey (Blum, 1965), and Greece (Williams, 1965). Evolution of the results of the use of this model were reported by Tinbergen and Bos (1965b) and Hollister (1967). The developers of this procedure conducted a few visual inspections of available data and concluded that discrepancies did exist and were likely to be a function of the fact that equation coefficients were not too accurate and the the linearity assumption regarding the parameters was likely to be untenable.

Hollister conducted an extensive analysis of the results and concluded that the model was indeed responsive to manpower demand, but improvement in manpower forecasting was needed. He concluded that the model should allow for the "substitution" of different educational levels, as well as level of skill as related to productivity, since the relationship of educational levels within an occupation drastically affects the stocks of educated manpower required to attain a specific level of economic activity.

Johnstone (1974) and Philp and Johnstone (1974) applied the model to Australia for the years 1954 and 1966. Little agreement between the Tinbergen-Bos model results and actual stocks of educated manpower was found. The authors felt this failure of the model to predict enrollment for the Australian secondary school system was particularly noteworthy, given the extent of economic growth that had occurred during this time period in that country. Any iterative model such as the Tinbergen-Bos model will be subject to increasingly discrepant results to the extent that the predictions used as input for future predictions are in error and the errors obtained do not cancel out as further iterations are completed.

In an extension of the Bolt et al. model, Reisman (1966) developed a model to encompass the four higher education production sectors--undergraduate, master's, doctoral, and post-doctoral level degrees. Reisman utilized the social systems simulation methodology developed by J. W. Forrester to solve differential equations interrelating system elements in the various equations. This modeling procedure produced estimates of the numbers of graduates expected within the four defined sectors of higher education.

This methodological approach was further developed by Reisman and Taft (1969) to include the flow of foreign nationals who study and work in the educational and noneducational sectors of the United States. Included within the extended model were provisions for the psychological, sociological, and economic factors which were thought to influence the movement of personnel within the educational and non-educational sectors. Over 200 nonlinear difference equations were used to simulate the production of university degree-holders and their feedback into higher education. Difference equations were utilized in this model because of the limitations of the Forrester processor, even though it was assumed that continuity of student flow was more realistic.

Hammond (1968) proposed a model to depict flow among undergraduate, graduate, university faculty, and professional nonacademic engineers and scientists. Four nonlinear equations were developed which described the number of individuals within the five groups as a function of an independently determined growth rate and the number of full-time faculty with a Ph.D. or equivalent degree. Hammond assumed, like Bolt et al., that the parameters of the model remain constant or change very slowly over time. Utilizing information available on numbers of individual in in the five educational/professional categories in 1961, Hammond developed estimates for the parameters of the model. These coefficients were used to generate projections for several years into the future. Hammond suggested that the validity of this model was demonstrated by the agreement of estimates and predictions developed independently by professional agencies for the year 1970.

In 1966, Correa presented a system model of the elementary and secondary educational structure based on his work for OECD. The model utilized information on the classroom periods of education offered in a day and on the number of periods of education in the particular educational subsystem being studied in order to generate enrollment estimates at a particular level of aggregation. While the model was based upon the Tinbergen-Bos approach, no information is provided which can be used to determine the validity of the proposed model to represent student flow.

This completes the review of a general class of simultaneous equation models operating without constraints on inputs or outputs. All of the above models cited use the difference (differential) equation approach to represent change over time of enrollment. One particular class of siumultaneous equation models is the Markov Chain model. As will be noted later, when the assumptions underlying the Markov model are relaxed, the resulting matrix equation can be rewritten as a set of linear, first-order difference equations. The presentation of Markov Chain theory will be followed by applications of Markov Chain models for purposes of representing student flow.

Markov Chain Models

The Markov Chain models are based upon the Markov Chain theory which was first presented in 1912. The basic assumption underlying the theory is that transition of an element (student) from one state (e.g., educational grade or course) within the model to another mutually exclusive state during a discrete interval of time is dependent only upon the present state of the element and is independent of any activities in the past. The formulation of student flow using the Markov model theorems is presented below.

A matrix P is defined as:

$$P = \begin{matrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1k} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2k} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3k} \\ p_{k1} & p_{k2} & p_{k3} & \cdots & p_{kk} \end{matrix}$$

where P_{ij} represents the probability that a student will move from one educational state at time t to another educational state at time t + 1. Parameters of the model are estimated by the obtained transition proportions of individuals in state i at time t who will be in state j at time t + 1.

By definition, the transition matrix elements P_{ij} are non-negative (i.e., probabilities cannot be less than 0.0) and the rows must sum to unity ($\sum_{j=1}^{k} P_{ij} = 1$) so as to preserve the attribute of mutual exclusiveness of system states. Time intervals may be defined as quarters, semesters, or academic or calendar years. (The term state as used in this paper refers to settings in an educational system where students receive organized instruction.)

Assuming that the distribution of the population, represented as a vector Y, among the educational states is known for an initial time period t, then the transition matrix can be used to predict the relative frequency of students who should occur among the educational states after a time interval (t to t + j). The Markov-type model assumes that from time t to time t+1, a single student either remains in his original state (i.e., is recycled back to where he/she started) or moves to one, and only one, other state.

This model can be compactly written according to matrix notation. Equation (15) expresses the Markov Chain model.

$$\underline{Y}_{t+1} = \underline{Y}'_t P \tag{15}$$

where

(a) $\underline{Y}'_t$ is a K x 1 row vector of the original distribution of students in K states at time t

(b) P is a K x K matrix of transition probabilities of student movement from state i to state i or j at time t+1

(c) $\underline{Y}_{t+1}$ is a K x 1 column vector of the distribution of students in K states at time t+1

Thus, the matrix multiplication of the transpose of input vector $\underline{Y}_t$ by the transition matrix P will give the expected output vector $\underline{Y}_{t+1}$. Thus, as new input entries and transition probability estimates are made available for a sequence of time periods, each of equal length, the population in each of the categories for each of the succeeding time periods may be estimated on the basis of the above matrix multiplication. It is also possible to use predicted output for one year as input to another iteration of the Markov Chain model. Markov Chain models are, by theory, stochastic in nature. The mathematical properties are that the collection of probabilities represented by the transition matrix will remain constant over time (the stationarity assumption) and that the probabilities are independent of any prior status of the individual no step memory assumption.

Following the usual approach to applying Markov Chain theory to student flow modeling, the element of the empirical transition matrices are the numbers of students (n_{ij}) moving from one educational state to another state or remaining in the same state. The adjustment of these transition frequencies for the total number of subjects observed in the total system gives an estimate (i.e., a transition frequencies for the total number of subjects observed in the total system gives an estimate (i.e., a transition proportion (p_{ij}) of the transition probability presented in equation (16):

$$p_{ij} = \frac{n_{ij}}{n_i} \quad . \tag{16}$$

Any single transition proportion p_{ij} can, thus, be seen to represent a particular ratio value much in the sense of a survival cohort ratio. If one uses p_{ij}'s as in equation (15) without being able to demonstrate the validity of the stationarity and no-step memory, then equation (15) represents a simultaneous calculation of K cohort survival ratios for original state i. Since cohort-ratio calculation methods are deterministic in nature, it is obvious that such a Markov Chain model must also be

considered deterministic. Economists have also pointed out the similarity between this approach to enrollment forecasting and the input-output type of economic analysis developed by Leontiff.

Most applications of Markov Chain theory models can be further categorized by whether they represent absorbing Markov Chains or ergodic Markov Chains. An absorbing chain has at least one model state which completely "absorbs" the element so that movement out of this state is impossible. Graduation from elementary school is one example of an educational state which could be included in an absorbing chain model. Ergodic chains allow movement between all states in the educational system being modeled.

Examples of Markov Chain Models

The attractiveness of the Markov Chain for modeling student flow through an educational system appears to have first been considered by Brown and Savage (1960). In this part of the Minnesota Methodological Studies on Enrollment Forecasting, the researchers were concerned with attempting to forecast the number of students enrolled in the colleges (i.e., academic division) by class, year, and sex.

As part of their project to test different methods for estimating future enrollments, registration records of students at the University of Minnesota were matched for a two-year period, 1957-1959, to give the actual numbers of students who moved into, stayed in, moved between, or moved out of various curricula. (No distinction was made for full- or part-time students, and no attempt was made to adjust for dropouts.)

Separate transition matrices were calculated for each sex to describe the student flow between colleges and academic standing (i.e., freshman, sophomore, junior, senior, graduate, or extension). The transition matrices were utilized together with the projections of the enrolling freshmen obtained in other phases of the project to generate the estimates of the numbers of students likely to be in attendance in the various colleges for the fall quarters of 1960 and 1961. The expected enrollments for time t were based upon the multiple linear regression models discussed earlier in this report.

The resultant enrollment forecasts were based on a single transition matrix for each year and sex combination, thus providing four sets of output. While no estimates could be obtained concerning the stability of the observed transition probability estimates, a visual comparison of the obtained predictions against the actual fall, 1960, enrollment by curriculum showed fair agreement. (A visual inspection of the transition matrix also indicated that the underlying assumption of Markov Chain model theory would not be met.) These results were considered by Brown and Savage to be interesting in view of the lack of research on the use of Markov Chain models in forecasting enrollment.

The Markov Chain model has been used in a variety of educational planning situations where a need for estimates of student enrollment exists. Thus, the remainder of this discussion of Markov Chain Theory will present examples of where it has been used for (1) single institutions (i.e., colleges or universities) and (2) regional or national educational systems. Distinctions between the various applications will be made where appropriate.

Single institutional Markov Chain models. This method has found wide use at the higher educational institution level in the United States, probably because of the original research provided by Brown and Savage at the University of Minnesota (1960). Other four-year institutional models have been developed for Michigan State University (Gani, 1965), Columbia University (Wong, 1969), the University of California at Berkeley (Marshall, Oliver, and Suslow, 1970), Rensselaer Polytechnic Institute (Bosuck and Wallace, 1970), and Illinois State University (Harden and Tcheng, 1971). While the Michigan State model was a straightforward application of the Markov Chain model to the numbers of students observed at the various educational levels within the university iterated by quarter, the other models did include some interesting variations. The Columbia University model was designed to have a stochastic basis, since transition probabilities entered within the transition matrix were sampled from probability distributions based upon observed data obtained for 40,000 students collected over a period of five years.

The Rensselaer model is of interest because the researchers utilized regression analysis procedures to estimate future transition probabilities between states given previous transition probabilities for those same states. Data on students were collected such that each student could be categorized by academic year, school within the institution (i.e., Architecture, Engineering, Humanities, Management, and Science), and four demographic factors of Rensselaer students (i.e., sex, age, geographic origin, and annual credit hour load). Using the transition matrix obtained by sampling students according to the identified demographic factors and then by iterating the model, comparison of actual and predicted data (from the simulation) for the years 1963 through 1967 showed a range of mean difference of -1.5 to -11.2 percent. However, an inspection of the summary of discrepancy results showed a much poorer fit for the last two years than for the first two. Use of this model on total New York statewide system data showed projections differing from those developed by the Office of Higher Education Planning, New York State Department of Education, by, at most, three percent. While these results are encouraging, one is led to wonder on what basis the New York State projections were generated. If the same faulty assumptions were used in the Rennselaer and State Department models, then one would expect agreement in model outputs and disagreement between actual and predicted results.

The Illinois State model investigated the use of Markov Chain models where enrollment ceilings existed. Data presented showed relatively good agreement at lower academic levels of the institution; but as movement between programs increased at the upper levels, transition proportions seem more variable with a resultant loss in accuracy.

A study of the effectiveness of four alternative approaches for developing enrollment projections, including the Markov Chain model, was performed by the American College Testing Program at the Kansas State University (Orwig, Jones, and Lenning, 1972). Using freshman enrollment figures for a five-year period beginning in 1965 for ten different departments, the baseline, trend line (single regression), simple ratio (cohort-survival), and Markov model methods were used to estimate enrollment for the separate departments. A final institutional total was determined by summing individual department totals. In order of overall effectiveness, the baseline method, a simple ratio of departmental enrollment to total institutional population, was most effective, followed by simple trend, Markov Chain, and simple ratio. However, the Markov Chain was relatively more effective in categorizing students than the baseline method in five of the ten prediction situations considered.

Regional or national educational system Markov Chain models. The use of Markov Chain models at the regional or national educational system level has also been quite popular. These highly aggregated models would seem to have the advantage of providing more stable estimates of student transition by the very fact that estimates are based upon larger sample sizes.

Models on the regional level have been developed to predict enrollment for all education (i.e., primary through higher education) in the State of New South Wales in Australia by Johnstone and Philp (1973) and for postsecondary educational enrollment in the State of Washington (Washington State Higher Education Facilities Commission, 1970), the State University System of New York (Bosuck and Wallace, 1970), and the province of Ontario, Canada (McReynolds, 1971). The work of Johnstone and Philp is interesting because they have compared actual enrollment to predicted enrollment obtained for the period 1947-1961. A pattern of overestimation for the first half of the period and underestimation for the second half was found. It seems apparent that social demand mechanisms were operating to cause the observed deviations between predicted and observed results. It seems that the regression approach followed by Bosuck and Wallace (1970) in the Rennsselaer study would have been applicable in the Johnstone and Philp study. The Washington State model was noteworthy in that it developed an interinstitutional as well as intrainstitutional model for use in long-range planning for public senior high school education in the state. A similar model was devised for use in the North Carolina Community College System by Wasik (1971a) which provided for enrollment projections at three levels--individual institution, regionwide and statewide.

One of the features of McReynolds' research was the method of synthetic cohorts developed to provide an estimation of transition probabilities using only stock data. Utilizing the synthetic transition matrices along with derived functions of student ability, McReynolds provided estimates of the numbers of students likely to enroll in programs aggregated within institution type. Fair agreement was found for these approaches to representing student flow.

Markov chain-type national educational system models have been developed for Jordan elementary schools (Wederlin, 1966), England (Moser and Redfern, 1965; Armitage, Smith, and Alper, 1969; and Stone, 1965, 1966, 1972), the Netherlands (OECD, 1967), and the United States (Zabrowski, 1969; Li, 1970).

The work of Thonstad (1967, 1969) is generally recognized as one of the most sophisticated attempts to use Markov Chain theory to represent flow of students through a system. The purpose of the model is to examine the long-range implications of the existing educational policies for producing various types of educated manpower. Given an input of flow of students between levels, the model produced numbers of students to be expected in each educational state for a specified number of years, the average number of years of education remaining for the student in a system state, and the percentage of students finishing each level of the educational system. Sixty educational states were identified and used in the modelling process. The model gave relatively good predictions for the primary and secondary levels of the Norwegian educational system, which are compulsory, but it was much less accurate when developing predictions for the higher education sector. Thonstad commented that the resulting inaccuracies at the higher education level were functions of the fact that there was limited space available in this part of the system. Thus, a drawback of the model is that it is least accurate where the educational system must be totally responsive to social demand, a drawback which is likely to lessen the usefulness of this type of model for the typical educational system.

A comprehensive model of the English national educational system is described by Armitage, Smith, and Alper (1969). A basic Markov Chain-type flow model serves to represent flow of students between basic levels of education and the feedback of students into the educational system as teachers, or the movement of students into the absorbing state of the "outside world." The model is then modified to accept the possibility of "bottlenecks" where capacity of a particular educational state is less than the students demanding entrance.[6] Thus, this model takes into account one of the facts that may have mediated against the ability of the Thonstad model to represent accurately student flow.

---

[6]The problem of bottlenecks is described as almost solely occurring in the higher education sector of the British educational system.

A rational model for deciding where the overflow of students from the bottleneck in the system will go is also treated. Some preliminary work in using this model for investigating policy questions by simulation is also presented, while further work reported by Armitage, Phillips, and Davies (1970) does shed some light on the form of the flows of students moving from secondary to higher education.

Stone's work (1965, 1966, 1967) has been based upon the use of Markov-type models in trying to represent individual person flow between levels within a particular social system. His development of a Markov model approach moved from the original model (1965), which proposed an epidemic-type Markov model for representing movement between levels of industrial and professional training as well as education received in educational institutions, to a general approach for representing a model of general approach for representing a model of general social system behavior (1972). Stone believed his contribution to model development to be unique in that he showed how two accounting systems--demographic and economic--can be interrelated so as to form what could be considered an input-output equation. Thus, an advantage of his approach is the determination of input-output coefficients to translate final educational demand into total output. However, much of the work presented is designed to describe the movement of individuals within the social system of interest, not necessarily to allow decision-makers a tool for use in policy choice.

Personnel at the Division of Operations Analysis, National Center for Educational Statistics, U. S. Office of Education, have also been interested in the use of Markov Chain approaches to modeling student flow. Zabrowski (1969) developed a computerized Markov Chain-type demographic flow model, which was subsequently named DYNAMOD II, to compute the numbers of individuals in 140 distinct demographic groups over selected periods of time. Separate transition matrices were estimated for each of four sex-race groups across 30 age-educational states, with the absorbing state representing death. Population inputs were obtained from the 1960 Bureau of the Census 1/1000 sample data tape. Estimates of numbers of births were also obtained from the Bureau of the Census.

DYNAMOD II is a modification of the Markov Chain model in that a forcing function of the distribution of births is included as a separate and variable input at the end of each time interval under consideration. The model equation states that the number of persons in a particular state in the year t+1 is equal to the number of persons in that category in year t who remained plus the number of persons who switched to that state plus births in the appropriate state. Zabrowski found a reasonably close match of data provided by DYNAMOD II and the ten-year projections developed by the Office of Education and the Bureau of the Census as one attempt to validate the model.

Three separate simulation experiments were conducted to demonstrate the feasibility of using a Markov Chain flow model (e.g., DYNAMOD II) to test the effects of alternative administrative policy decisions. The experiments were performed in order to determine the effects of policies

designed to (1) increase retention rates of students in the elementary, secondary, and college levels in the United States, (2) change the student-teacher ratio if retention rates were raised or lowered to specified percentage levels, and (3) determine an optimal sequence to be followed by administrators in implementing policies to affect retention rates.

Another demographic model designed to model student flow was presented by Li (1971). In his attempt to use Markov Chain theory to represent the dynamics of student cohort progression, Li considered mortality and dropouts to have less of an influence on student population than repetition and new student entries. He presented procedures for estimating grade-specific droupout and martality rates from age-specific data. Grade-specific repetition rates were solved for directly by use of a recursive equation.

Optimizing Simultaneous Equation Models

A number of authors (e.g., Correa, 1967a; Hammond, 1970; McNamara 1971b) have argued for the use of optimization of solutions provided by a mathematical model. These recommendations are based on the belief that resources available to an educational system are finite and subject to certain constraints.[7] A number of procedures are available for use in providing such optimal solutions, and they are of the type generally utilized in operations research activities.

The general idea behind the use of constraints in a mathematical model will be presented here. Assume that a specific output y is to be estimated from a knowledge of independent variables x_1, x_2, . . . xp by a multiple linear regression equation, where some of the k independent variables are subject to constraints. The resultant model is a linear (in the parameters) programming model. The equations used to provide a solution would have the following general form:

$$(17a) \quad y = b_1x_1 + b_2x_2 + b_3x_3 + \ldots b_px_p$$

$$C_1 > c_{11}x_1 + c_{12}x_2 + c_{13}x_3 + \ldots c_{1p}x_p$$

$$C_2 > c_{21}x_1 + c_{22}x_2 + c_{23}x_3 + \ldots c_{2p}x_p$$

---

[7]Use of extrapolication or simultaneous equations procedures to predict the future in an expanding enrollment base year can result in the education sector's accounting for all available social service funds after only a few iterations. Since this is most unlikely to happen in practice, the use of constraints on resources available for use in controlling systems seems to be entirely reasonable.

(17b) .

.

.

$$C_n > c_{n1}x_1 + c_{n2}x_2 + c_{n3}x_3 + \ldots c_{np}x_p$$

(17c) $x_j > 0,\ j = 1,2, \ldots k(k < p).$

Equation (17a), the "objective function," has the form of a multiple linear regression equation and is used for purposes of prediction. The equations in (17b) represent the constraints which can be imposed upon the independent variables represented in equation (17a). The last equation (17c) represents the constraints to be imposed upon the values of the independent variables of the prediction equation (i.e., the independent variables can only take on positive values). The total set of equations (17a, 17b, and 17c) is simultaneously solved so that a solution for the maximum (or minimum) value for y is obtained that allows for the independent variables' values to vary within the prescribed limits imposed upon the independent variables.

Operations research methodology has been applied at higher education levels to assist college administrators in resource allocation (Flessner, Fox, and Sanyal, 1968; Crandell, 1969; Mann, 1970) and in the allocation of faculty (Fox, 1963; Andrew, 1970). The majority of these studies have used linear programming procedures.

The use of mathematical programming systems at national and/or regional (i.e., state) levels can provide the greatest amount of payoff because of economies of scale. For this reason, economists have been quite interested in developing large, relatively complex models which include student flow as one component and which attempt to aim an economic system in the direction which would support national goals of economic development. These models have been developed from the philosophic basis implied by the manpower demand approach to educational planning; that is, orderly economic growth can occur only if trained manpower is available for maintenance and expansion of a nation's industrial base.

One of the most ambitious attempts to use linear programming procedures has been presented by Bowles (1965, 1967). This particular model used the linear programming methodology to determine what amount of Nigeria's resources should be devoted to educational development and how such resources should be allocated within the educational subsystem so as to maintain the flow of trained manpower needed to attain desired economic growth goals. This model considers all students at all levels of the educational system and the labor force of Nigeria. Outputs of the model include cost relations of the proposed educational system to the national economy and the expected economic benefits accruing from the educational system to the economy, as well as the optimal flow of students between educational system levels needed to meet the

desired economic goals. The model requires, as input, student flow data in the form of input-output matrices (i.e., transitions between levels of the educational system represented as proportions), the cost factors of the educational system, and estimates of income by educational level. Thus, this model provides for specification of constraints of resources which are applied to the expansion of the educational system.

Other national models have utilized similar procedures have been developed by Adelman (1966) for Argentina and By Bernard (1967) for the French educational system. A model developed by McNamara (1971a) is similar in spirit though slightly more limited since it is concerned with only the flow of individuals from vocational education subsystems into the work force. Also, no total level of economic production is considered.

The Harvard University Center for Studies in Educational Development was concerned with the development and evaluation of alternative mathematical programming procedures for use in educational planning in the late 1960's (Davis, 1968). Among the model forms considered and then discarded were static, quasi-dyanmic and pure dynamic programming, quadratic programming, control theory, and analogue simulation. It was concluded that no single model form, be it optimizing simulation or otherwise, would generate the necessary information for an educational planner to make accurate forecasts.

Finally, note is made of the claims of Alper (1967, 1971) and Tracz (1971) that use of control theory concepts is appropriate in the solution of educational planning problems. Alper and Smith (1967a, b) suggested that these procedures are especially useful in considering bottlenecks which result from greater demand for places in a particular state of an educational system than can be accommodated at a particular point in time. While a model is provided in Alper and Smith (1967b), there are no data available to provide an evaluation of the effectiveness of control theory approaches except for the results presented by Davis, which were not encouraging.

APPROPRIATENESS OF MATHEMATICAL MODELS FOR STUDENT FLOW

The preceding discussion has presented a general outline of methodologies utilized for representing student flow within educational systems. It has been noted that some mathematical computational approaches have achieved a great deal of popularity for developing student flow models. It has also been previously suggested that this popularity may be more a function of the ease of application than of the appropriateness for representation of the system of interest.

An evaluation of the appropriateness of a particular mathematical model for representing the student flow in an educational system should be based on objective criteria. Four criteria seem to be especially relevant to the question of determining the appropriateness of a mathematical model. These criteria will now be presented and discussed in turn.

Underlying Assumptions

As noted earlier, the most popular approach to modeling student flow through educational systems has been through the application of Markov Chain theory. However, Markov Chain theory posits two assumptions which must be met if the theorems associated with mathematical theory are to be applied. These are the assumptions of one step memory and stationarity. While large sample statistical tests (when n is large) are available for determining whether a particular transition matrix will meet the assumptions underlying Markov Chain theory, this reviewer has not been able to find any model based upon Markov Chain theory which did test the underlying assumptions. Moreover, neither of these assumptions seems reasonable in the case of using Markov Chain theory to represent student flow. It would appear that many individuals who apply the Markov Chain theory to educational planning problems subscribe to the strong statistics notion that arithmetic manipulation of numbers can be performed independent of the meaningfulness of the final product. Creager (1970) presented a case for utilizing the matrix multiplication associated with Markov Chain theory without requiring the assumptions to be met as a way of thinking about the representation of masses of data collected on student flow.

If the researcher is only interested in description, as was Creager, then it seems reasonable to disregard the theory requirement. However, if one is interested in the prediction of some future event, it would seem prudent that some consideration be paid to the extent to which available data fit the Markov model.

Multiple regression is another method that has been used extensively to estimate enrollments for a variety of educational systems. Concerns relating to regression analysis assumptions deal with whether the sample represents an independent sample and also whether the additive

nature regarding the independent variables is tenable. In the case of non-independent sampling of data (i.e., time series data), the use of trands to fit the data provided by enrollment period can be misleading if the sample size is not large. However, even with a large n, the data published in the Minnesota educational attendance study tend to suggest that exogenous variables in addition to time must be used to obtain accurate estimates of enrollment. Impara (undated) found that some grade enrollments for the State of Florida were best predicted by a polynomia form of a regression equation; however, no rational explanation could be developed which would explain the necessity of curvilinear terms in the model. It is also notable to this observer that no attempts to use cross sample validation formulas which correct for shrinkage have been reported in the use of multiple regression procedures to model student flow. In situations where the sample is small, drastic crops in predictive efficiency would be expected when attempts are made to use the equation for prediction in a new situation. A linear programming algorithm has been the procedure used in the majority of optimizing mathematical models discussed previously in thsi review. These tendencies have persisted even though researchers (for example, Davis, 1968, and Sisson, 1969) with substantial experience in the application of operations research methodology to educational planning models building have argued for nonlinear representation of relationships of model variables. Therefore, it seems likely that linear programming models have achieved popularity because of ease of application and availability of computer routines to perform required calculations, and not because of the theoretical desirability of the linear relationship assumption.

It is noteworthy that the assumptions concern is raised only with reference to models developed within a stochastic framework. Deterministic models do not require specific assumptions regarding the form of data used in the model. Thus, even though the theory assumes a stochastic data structure, the actual use of the Markov Chain and regression models for enrollment estimation seems to follow a deterministic philosophy.

Validation of the Model

The question of how well a model mimics and/or predicts reality is of primary importance in any model building exercise. As noted earlier, a primary consideration of the validation phase of a study revolves about determination of whether the stated purpose of the model is achieved.

While a model which is designed as an analogue to the system of interest does not have to predict perfectly to be useful, it is rather difficult to see how a model can mimic exactly the behavior of a system and not provide accurate predictions. A poorly predicting model must stand accused of not having considered all variables which influence the behavior of the system.

In contrast, the model designed to predict future events does not necessarily mimic exactly the behavior of the system of interest, but must provide accurate estimates. Thus, it is not surprising that a different mathematical procedure can be proposed and found to be useful in the estimation of enrollment in various educational activities within a system. However, the question of determining how accurate a model's projection must be before the model can be considered validated does not avail itself of a clear-cut type of answer. A response, "accurate as possible," may not be a vaiable answer if one considers the fact that accuracy of model estimates is likely to be associated with the detail of the model which, in turn, is a function of the resources expended on the development of the model. Thus, the cost of developing a model to provide estimates which have one-half the error of prediction of a second model will likely cost much more than twice the cost of the second model. That is, as one approaches a level of predictive accuracy, as yet unknown, the cost of accomplishing the desired level may increase to a level that could not be supported on the basis of economic benefit realized from the use of the more accurate model.

A second consideration regarding the development of accurate pro-projections is that the proposed use of the model may achieve a higher degree of accuracy than that required by the procedures which use the model's output. Thus, resources may be wasted in the development of complex models when relatively simpler models can prove to be just as effective in providing the information needed (or can be used) by decision-makers. The use of a simple trend line regression approach with a single linear parameter in time may provide enrollment projections of the level of accuracy that can be handled by the educational system manager, even though a more complex regression model with trend components and exogenous variables might provide a more accurate estimate but at a higher cost of data collection and other resource utilization.

One final concern relates to how one can determine if a model is sufficiently accurate, whatever the degree of error associated with the model output. A problem arises from the fact that model outputs do not represent independent observations, and, thus, the general statistical methods available for use with frequency-type data are not, technically speaking, appropriate. Creager (1970) has suggested a chi square test for use in comparing outputs from a Markov Chain model, although he recognizes the above objection of the lack of independence of data as required by the test he proposes. It seems that if an individual is willing to overlook the assumption of independence, then chi square tests may also be used with caution. However, this lack of statistical tests which can be used in validating models does lead to a reliance on subjective approaches such as errors of prediction, reasonableness of predictions, etc.

Availability of Data

Lack of data and/or inappropriateness of the format of available data have been identified as major impediments to the development of a mathematical model of the educational system and subsystems by model builders. As long as educational data are collected by organizations which are not in communication with model builders, as is typically the case now, this complaint is likely to be recurrent.

There seems to be at least four obvious remedies for this problem. The first, and most straightforward, would be for the model builder to go to the individuals (or organizations responsible for collecting the data and ask that the data-collection peration be changed so that data can be summarized and made available in the form which could be used directly in the mathematical model. While this may be a rational approach to solving this problem, it would seem likely that unless the model builder has the support of the highest level of management, the individuals with responsibility for collecting the data will tend to ignore the request. Further, the possibility of the data-reporting agency's changing its procedures would seem to be an inverse function of the degree of change required within the data-collection and/or data-reporting reporting system. This is understandable since agency resources would have to be committed to achieve the necessary operational changes, and the agency's manager would have to consider the appropriateness of the use of such resources for other organizational needs.

A second possible solution to this problem is to recognize the limitations of the available data and utilize a mathematical approach that would require only data collected in the present form. This suggestion may be criticized as showing opportunism and ignoring the advantages of using mathematics to develop a model for a particular situation. However, such an approach to model building is no more inappropriate than the attempts to use certain mathematical methods in modeling because these methods are easy to apply.

A third approach is to change either the form or the type of variables used to represent the model if appropriate data are available for one set of variables but not for the variables required in the original model. An example of this approach was provided by Stone (1968) who changed his model structure so that the variables in his model would correspond to the variable definitions provided by the United Nations System of National Accounts. Stone admitted the process of changeover from the original model variable formulation to the U. N. system was not only costly in terms of resources expended, but it also brought research to a standstill because of the time required to relabel and revalue the model variables. This expenditure of resources was judged well spent, however, because of the increased accuracy of the data obtained from the U. N. data system and because information would be available on variables not considered in the original model.

The fourth approach is to transform the available data into a format that is appropriate for use with the particular mathematical model of interest. Such an approach was followed by McReynolds (1971) who developed a synthetic cohort method for utilizing stock data to estimate the transition probabilities of individual students moving from one educational system state to another. Data available from the Ontario Carnegie Study of Identification and Utilization of Talent in High School and College on the careers of approximately 90,000 students who were in grade 9 of the Ontario, Canada, schools in September, 1959, were used to generate synthetic transition proportions. These proportions showed relatively good fit to the actual transition proportion information available, thus providing support for this particular method of using stock data to estimate transition proportions.

In general, educational statistical data tend to be aggregated; subject, etc. An aggregated data system, even when classifications are not highly aggregated, provides information on distributions of students among various educational activities. However, the educational stocks cannot provide direct measures of the flow of students between the various educational system states since individuals are not followed, but rather, only the total is obtained. Thus, the number in our classification may not be only graduates from the next lowest grade or previous course; it could represent repeaters, migrants, in, or students coming from another track (i.e., students "double promoted" or students accepted into a course without having the normal course prerequisites). While these types of data would appear to be appropriate for use in the development of simultaneous equation models and extrapolation models, they are going to prove insufficient for the information needs for the estimation of transition probabilities required for the development of Markov Chain-type models.

In conclusion, it must be noted that there are a number of reasons why well ordered data should be useful in the development of mathematical models. First of all, the accuracy of any derived mathematical model will be constained by the availability of "good" data. A second major concern relates to the cost of developing estimates of model parameters. These costs, as noted in the comments about the validity of mathematical models, will be directly proportional to the accessibility and arrangement of the available data. These problems are particularly important when one attempts to develop a Markov Chain-type model.

Completeness of Model

A Fourth concern is related to the degree to which the model of interest has accounted for the variables which influence the dependent variable (e.g., student flow). The majority of the models presently being used to generate estimates of student flow are either regression based (i.e., static type models) or Markov Chain-type (i.e., dynamic type models). These models can be described as essentially main effects models in a statistical design sense. That is, the regression and Markov-type models assume implicitly that no interaction exists between levels of the various independent variables used to generate estimates of

of enrollment in the various educational system states considered by the model.

Smith (1970) investigated the effect of exogenous variables on predicting student movement between academic years, student movement between curriculum areas, and returning to school. He found that patterns of student flow differed between departments, which suggested that no single projection procedure would be applicable for each department or major. These findings could also be interpreted as providing support for the thesis that not all of the independent variables had been identified and included within the prediction model.

In a study of the flow of students moving through the English educational system, Orr (1971) was able to discriminate between subgroups of English school children who were likely to advance to the next educational level, move laterally to other types of study, or drop out of formal education entirely, by consideration of the interaction of various educational background and socioeconomic variables. This type of phenomenon is also well known by psychometricians as the moderator variable effect (Saunders, 1973). In the typical psychological testing situation, this phenomenon would be represented by the existence of different magnitudes of the correlation between two variables (one independent and one dependent) at differing levels of a third, independent variable. The predominant pattern in the Lea study showed attendance at the next highest level of education to be highly related to (1) whether the student attended an English public (i.e., read American private) or grammar (i.e., read American public) school and (2) the level of intellectual ability. Considering this finding in the moderator variable sense, the proportion of bright students going on to a higher level of education was higher from those who attended a public school than for those who attended a grammar school. However, the structural educational system factors accounted for proportionately less and the student background variables proportionately more of the total variation of student progress variables due to the high degree of selection of students resulting at the highest levels of the English educational system.

A second form of model incompleteness would be the failure to include all variables that would likely add to efficiency of prediction. Thus, even though interactions might not be considered, it is possible to increase the accuracy of model estimates by including variables which show high correlations with the dependent variable but are essentially uncorrelated or weakly correlated with other independent variables in the model. McReynolds developed three Markof-type models which utilized information on student academic ability, performance, or interest in addition to the information previously available on the patterns of student flow. While he was not able to test his ideas directly, it seems reasonable to conclude that his model would likely yield superior levels of predictive efficiency than Markov-type models do which use

only the knowledge of student status immediately prior to the change in educational state. It can also be noted that some Markov-type models have attempted to take into account the possibility of interaction of variables by disaggregating the levels of major independent variables. One particularly good example of this approach is provided by the University of Minnesota Markov transition ratio type of flow models. Individuals were classified not only by curriculum major but also by year in school (i.e., freshman, sophomore, etc.) so as to provide a more detailed look at the pattern of movement from one educational state to another over an academic year. This could be considered as one example of attempting to account for an interaction among the variables of curriculum major and year in college. The goodness of fit between prediction and actual observed results from this model has already been commented upon.

In conclusion, there are several difficulties associated with the development of mathematical models. Questions concerning the appropriateness of specific models for the particular situation in which they are to be used were considered in the previous four sections, along with suggestions for possible ways to overcome the stated concerns. At this point it seems appropriate to consider the future prospects of the use of mathematical models in educational planning, at least from the perspective of suggesting that there is a need for further support of the development of educational system models.

To evaluate the usefulness of mathematical models in educational planning, one must determine the effect educational planning has had on educational policy-making. Educational planning information in the form of enrollment forecasts has been widely accepted by educational system managers as useful in developing future budgets, and as such it serves indirectly to influence educational policy. However, it has been within only the last 15 years that educational planning has moved toward attempts to model entire educational systems as either a single system or a component subsystem in a regional or national economic model. Much of this increase in the level of complexity of educational system models has resulted from the interest of economists who were attempting to develop rational planning systems for economic development. It has been largely through the economists' claims of economic benefits accruing from educational planning that government agencies have been willing to support the attempts to develop mathematical models of educational systems.

Perceptions of Educational Planning Effectiveness

It seems fair to say that educational planning as a process has achieved a high level of sophistication and that much of this advance has been due to the development of mathematical models to represent

various aspects of educational systems. Moreover, the future level of support for development of quantitatively based educational planning is likely to remain a function of the perceived effectiveness of these methods. Thus, it seems appropriate to consider how educational planning is viewed by educational policy-makers--especially those who have overall responsibility for setting the goals and direction of the educational system of interest. It can be assumed that the perceptions of these individuals are likely to influence the extent to which future development of quantitative methods for educational planning is supported; these individuals will be responsible for the degree of acceptance of the use of educational planning information for policy-making. It seems reasonable to conclude, then, that the use of educational planning methodology is much more likely to increase if the policy-maker indicates the use of such procedures to be approved policy.

Pierre Billecocq stated in the chairman's report to the 1970 Paris OECD Conference on Policies for Educational Growth that the high expectations voiced at the 1961 Washington Conference on Economic Growth and Investment for the use of educational planning methodology to assist nations achieve specific economic growth rates had not been realized. He observed that while economic growth targets set at the 1961 Conference had already been matched or exceeded, "the growth of educational systems from 1960 to 1970 was an unsystematic rather than planner process, a response to social demand rather than an adjustment to the needs of the various branches of the economy" (Billecocq, 1971, pp. 26-27).

Clark (1971), in a later presentation at the same 1970 Conference, referred to the then current use of educational planning as "incremental" rather than "allocative." In other words, educational system managers were making use of information provided by educational planning to support increases in operational budgets, as opposed to providing information which could be used rationally to support the funding of one set of educational system alternatives versus any other particular set.

Both of these commentators noted that expenditures for education had increased over the past decade at a higher rate than expenditures used to support other social services mainly on the "unchallenged assumption that education is 'good'--in any form and in all possible amounts--for economic, if no other reasons" (Clark, 1971, p. 122).

Thus, the above post-mortem suggests that while some individuals may propose such an economic basis undergirding the expansion of educational systems in the OECD countries, such a proposition cannot be evaluated for validity. In OECD countries, increases in educational system sizes went along with economic growth, even though rates of increases may have varied. While all governmental policy-makers believed economic expansion was due, in part, to the availability of trained manpower, there was no possible way of determining what part of the economic boom was the effect of increased support of the educational system.

Implications of the failure to utilize results available from educational planning efforts are now becoming clear. In the industrial OECD countries and the United States, educational system graduates are no longer guaranteed a position in the field in which they are trained. For example, the overproduction of physical scientists at the advanced university graduate levels has been a problem since 1970. Recent advance estimates indicate that approximately four percent of the 1975 liberal arts graduates have jobs waiting for them in the field in which they were trained. This situation generally results in a job-seeker's acceptance of a position which previously required less education than he/she obtained. Thus, attempts to generate manpower demand estimates over time will likely result in erroneous occupational-educational linkage requirements; i.e., occupational education requirements will be set higher because of the influx of well educated individuals into an occupation who displace less educated individuals. The implication is that some individuals will have overinvested in their educational training, with a resultant drop in the relative financial advantage which had been traditionally associated with the college degree. Using data available for the decade of the '60's, Nollen (Boffey, 1974) showed that college attendance of young white males was a function of economic costs and benefits of a college diploma. Nollen concluded that young white males do behave as if the "investment-return" from a college education is an important factor in their decision to attend college. In this case, it appears to be the effect of competition in the job market rather than any educational policy based on rational educational planning which is associated with the production of college educated white males.

One conclusion thus derived is that educational planning as a management information tool has had little apparent effect on changing educational policy. However, this seems to be more a function of the failure of the policy-maker to call upon the expertise of the educational planners. The power of educational planning procedures is derived from the information provided on the implications and consequences associated with alternative strategies designed to assist in meeting specific educational system goals. While educational policy-makers may in some cases be interested in obtaining data to use in making policy decisions, they usually do not know what data are needed or make unrealistic demands simply to meet congressional/parliamentary requests for information. Thus, until the educational planner is involved in the planning process which develops long-term educational goals or alternative educational policies, and until his views are seriously considered, the amount of measurable economic/social benefit resulting from the use of educational planning is likely to remain disappointingly poor.

LIST OF REFERENCES

Adair, C. D. "Predicting New Year's Enrollment." School and Community, LV (1969), 37.

Adelman, I. "A Linear Programming Model of Educational Planning: A Case Study of Argentina. In I. Adelman and E. Thorebeche (eds.), The Theory and Design of Economic Development. Baltimore: Johns Hopkins Press, 1966, 385-412

Alper, P. "Introduction of Control Concepts in Educational Planning Models." In Mathematical Models in Educational Planning. Paris: OECD, 1967, 259-273.

Alper, P. "Some Considerations of Planning Engendered by Modern Control Theory." Socio-Economic Planning Science, VI (1971), 295-304.

Alper, P., and C. S. Smith. "Control Strategies for an Educational Planning Problem." Journal of the Canadian Operational Research Society, V (1967), 159-174.

Alper, P., and C. S. Smith. "An Application of Control Theory to a Problem in Educational Planning." IEEE Transactions on Automatic Control, AC-12, II (1967), 176-178.

Andrew, G. M. Operations Analysis. Minneapolis: University of Minnesota, Department of Management Science, May, 1970.

Armitage, P. H., C. M. Phillips, and J. Davies. "Towards a Model of the Upper Secondary School System." Journal of the Royal Statistical Society, Series A, CXXXIII (1970), 166-205.

Armitage, P. H., C. S. Smith, and P. Alper. Decision Models for Educational Planning. London: The Penguin Press, 1969.

Ball, R. J. "Econometric Models." In Mathematical Model Building in Economics and Industry. New York: Hafner, 1968, 15-30.

Banks, J., and C. L. Hohenstein. Conceptual Models and Procedures for Predicting Higher Education Enrollment. Paper presented at the meeting of the Operations Research Society of America, Chicago, October, 1970.

Beach, W. A. Economic Models: An Exposition. New York: John Wiley and Sons, Inc., 1957.

Bernard, J. "General Optimization Model for the Economy and Education." In Mathematical Models in Educational Planning. Paris: OECD, 1967, 207-243.

Billecocq, P. "Address to the Conference on Policies for Educational Growth." In Educational Policies for the 1970's. Paris: OECD, 1971, 25-29.

Blum, S. "Planning Models for the Calculation of Educational Requirements for Economic Development - Turkey." In Econometric Models of Education: Some Applications. OECD Technical Report. Paris: OECD, 1965, 55-76.

Boffey, P. M. "A Degree's Value: Advantage Isn't as Great as It Used to Be." The Chronicle for Higher Education, IX:12 (December 9, 1974), 2.

Bolt, R. H., W. L. Koltun, and O. H. Levine. "Doctoral Feedback into Higher Education." Science. CXLVIII (1965), 918-928.

Bosuck, A., and W. A. Wallace. "A Computer Simulation Approach to Enrollment Projection in Higher Education." Socio-Economic Planning Science, IV (1970), 365-381.

Bowles, S. The Efficient Allocation of Resources in Education: A Planning Model with Applications to Northern Nigeria. Unpublished doctoral dissertation. Harvard University, 1965.

Bowles, S. "The Efficient Allocation of Resources in Education." Quarterly Journal of Economics, VIII (1967), 189-219.

Brown, B. W., and I. R. Savage. Methodological Studies in Educational Attendance Prediction. Minneapolis: University of Minnesota, Department of Statistics, 1960.

Cartter, A. M. "Scientific Manpower for 1970-1985." Science, CLXXII (1971), 132-140.

Clark, R. C. "Keynote Address to the Conference on Policies for Educational Growth." In Educational Policies for the 1970's. Paris: OECD, 1971, 119-124.

Correa, H. "More School and Better Schools?" Scientia Paedogogica Experimentalis, III (1966), 123-141.

Correa, H. "Models and Mathematics in Educational Planning." In G. Z. F. Bereday, J. A. Lawreys, and M. Blaug (eds.), The World Yearbook of Education, 1967: Educational Planning. London: Evans Brothers, 1967, 398-422.

Correa, H. "A Survey of Mathematical Models in Educational Planning." In Mathematical Models in Educational Planning. Paris: OECD, 1967, 21-93.

Crandell, R. H. "A Constrained Choice Model for Student Housing. Management Science, XVI (1969), 112-120.

Creager, J. A. Use of Empirical Transition Matrices in Educational Research. Paper read at the meeting of the American Educational Research Association, Minneapolis, March, 1970.

Cross, R. H., and C. H. Sederberg. "Computer-Assisted Enrollment Projection Procedures." Journal of Educational Data Processing, II (1969), 160-165.

Davis, R. G. "On the Development of Educational Planning Models at Harvard CSED: An Algebraic History of Activity in One Small Place." In R. H. P. Kraft (ed.)., Education and Economic Growth: Proceedings of the First Annual Conference on the Economics of Education. Tallahassee: The Florida State University, 1968, 42-84.

Draper, N. R., and H. Smith. Applied Regression Analysis. New York: John Wiley and Sons, Inc., 1966.

Emmerij, L. J. "Planning Models for the Calculation of Educational Requirements for Economic Development - Spain." In Econometric Models of Education: Some Applications. OECD Technical Report. Paris: OECD, 1965, 33-53.

Fattu, N. "Operations Research Applied to Education." Review of Educational Research, XXX (1960), 409-422.

Flessner, J., and K. A. Fox. On the Allocation of Resources in a Academic Department. Ames: Iowa State University, Department of Economics, 1967.

Flessner, J., K. A. Fox, and B. C. Sanyal. "On the Allocation of Resources in a University Department." Metroeconomica, XX (1968), 256-271.

Fox, K. A. Formulation of Management Science Models for Selected Problems of College Administration. Final report. U. S. Office of Education Contract No. OEC-3-068058. Ames: Iowa State University, 1963.

Gani, J. "Formulae for Projecting Enrollments and Degrees Awarded in Universities." Journal of the Royal Statistical Society, Series A, CXXVI (1963), 400-409.

Gani, J. A Model for Student Enrollment in American Universities: I. Structure of the Model. unpublished manuscript, 1965.

Gass, J. R. "The Evaluation of OECD's Approach to Educational Planning in Developed Countires." In G. Z. F. Bereday, J. A. Lawreys, and M. Blaug (eds.), The World Yearbook of Education, 1967: Educational Planning. London: Evans Brothers, 1967, 139-148.

Greenawalt, G. J., and D. P. Mitchell, Predicting School Enrollment. Cambridge, Massachusetts: New England School Development Council, 1966.

Haggstrom, G. W. On Analyzing and Predicting Enrollments and Costs in Higher Education. Paper presented at the meeting of the American Statistical Association, New York, August, 1969.

Hammond, A. A Flow Model for Higher Education. Santa Monica, California: Rand Corporation, 1968.

Hammond, A. Mathematical Models in Education and Training. Santa Monica, California: Rand Corporation, 1970.

Harden, W. R., and M. T. Tcheng. "Projection of Enrollment Distribution and Enrollment Ceilings by Markov Processes." Socio-Economic Planning Science, V (1971), 467-473.

Hollister, R. G. "A Technical Evaluation of OECD's Mediterranean Regional Project: Methods and Conclusions." In G. Z. F. Bereday, J. A. Lawreys, and M. Blaug (eds.), The World Yearbook of Education, 1967: Educational Planning. London: Evans Brothers, 1967, 161-169.

Impara, J. R. Long-Range Planning Model for Simulating Educational Manpower Needs. Unpublished manuscript. Florida State Department of Education, undated.

Jacoby, E. G. Methods of School Enrollment Projection. Educational Studies and Documents No. 32. Paris: UNESCO, 1959.

Johnstone, J. N. Mathematical Models in Educational Planning. Unpublished Master's thesis. University of Sydney, 1971. Cited in J. N. Johnstone, "Mathematical Models Developed for Use in Educational Planning: A Review." Review of Educational Research, XLIV (1974), 177-201.

Johnstone, J. N., and H. Philp. "Application of a Markov Chain Model in Educational Planning." Socio-Economic Planning Science, VII (1973), 283-294.

Kemeny, J. G., and J. L. Snell. Mathematical Models in the Social Sciences. New York: Ginn, 1962.

Kendall, M. G. "Introduction to Model Building and Its Problems." In Mathematical Model Building in Economics and Industry. New York: Hafner, 1968, 1-14.

Kershaw, J. A., and R. N. McKean. Systems Analysis and Education. Santa Monica, California: Rand Corporation, 1959.

Kuang, H. P. "Forecasting Future Enrollment by Curve-Fitting Techniques. Journal of Experimental Education, XXIII (1955), 271-274.

Li, W. L. "A Demographic Model of Student Progression." International Review of Education, XVII (1970), 408-424.

Lins, L. J. Methodology of Enrollment Projection for Colleges and Universities. Madison: University of Wisconsin, 1952.

Liu, B. A. Estimating Future School Enrollment in Devleoping Countries. New York: UNESCO, 1966.

Lucas, H. L. Personal Communication. June, 1974.

Maki, D. P., and M. Thompson. Mathematical Models and Applications. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973.

Mann, S. H. "Dynamic Management of Library Size." Operations Research, XVIII (1970), Supplement 2, B-149.

Marshall, K. T., F. M. Oliver, and S. S. Suslow. Undergraduate Enrollments and Attendance Patterns. Administrative Studies in Project in Higher Education, Report No. 4. Berkeley: University of California, 1970.

McNamara, J. F. "A Mathematical Programming Approach to State-Local Program Planning in Vocational Education." American Educational Research Journal, VIII (1971), 335-363.

McNamara, J. F. "Mathematical Programming Models in Educational Planning." Review of Educational Research, XLI (1971), 419-446.

McNamara, J. F. Modification of the Age-Cohort Method for Projecting First Grade Enrollment. Unpublished manuscript. University of Oregon, 1971.

McReynolds, W. P. A Mathematical Model for an Educational System. Occasional Paper No. 9. Toronto: The Ontario Institute for Studies in Education, 1971.

Moser, C. A., and P. Redfern. "Education and Manpower: Some Current Research." In C. M. Berners-Lee (ed.), Models for Decision. London: English University Press, 1965, 71-83.

Newton, R. M., and W. H. Thomas. "Design of School Bus Routes by Computer." Socio-Economic Planning Science, III (1969), 75-85.

Organization for Economic Cooperation and Development (OECD). Mathematical Models in Educational Planning. Paris: OECD, 1967.

Orr, L. "The Dependence of Transition Proportions in the Educational System on Observed Social Factors and School Characteristics." Journal of the Royal Statistical Society, Series A, CXXXV (1972), 74-95.

Orwig, M. D., P. K. Jones, and O. T. Lenning. Enrollment Projection Model for Institutional Planning. Research Report No. 48. Iowa City: American College Testing Program, 1972.

Philp, H., and J. N. Johnstone. "The Use of Mathematical Models as an Aid to Effective Planning in Education." In Planning for Effective Education. Carlton: Australian College of Education, 1971. Also cited in J. N. Johnstone, "Mathematical Models Developed for Use in Educational Planning: A Review." Review of Educational Research, XLIV (1974), 177-201.

Reisman, A. "Higher Education: A Population Flow Feedback Model." Science, CLIII (1966), 89-91.

Reisman, A., and M. I. Taft. "The Generation of Doctorates and Their Feedback into Higher Education." Socio-Economic Planning Sciences, II (1969), 473-486.

Saunders, D. R. "Moderator Variables in Prediction." Educational and Psychological Measurement, XXXVIII (1973), 209-222.

Sawhiris, M. Y. "The Projection of College Enrollment." Multivariate Behavioral Research, V (1970), 83-116.

Sisson, R. An Hypothetical Model of A School. Philadelphia: Fels Institute of Local and State Government, University of Pennsylvania, 1968. ED 030 978

Sisson, R. "Can We Model the Educational Process?" Socio-Economic Planning Science, II (1969), 109-119.

Smith, W. "A Student Flow Model." Mimeograph. Los Angeles: Office of Advance Planning, University of California at Los Angeles, 1970.

Stone, J. R. N. "A Model of the Educational System." Minerva, III (1965), 172-186.

Stone, J. R. N. "Input-Output and Demographic Accounting: A Tool for Educational Planning." Minerva, IV (1966), 365-380.

Stone, J. R. N. "Economic and Social Modelling." In Mathematical Models in Economics and Industry. New York: Hafner, 1968, 49-60.

Stone, J. R. N. "A Markovian Education Model and Other Examples Linking Social Behavior to the Economy." Journal of the Royal Statistical Society, Series A, CXXXV (1972), 511-532.

Thonstad, T. A. "A Mathematical Model of the Norwegian Educational System." In Mathematical Models in Educational Planning. Paris: OECD, 1967, 125-158.

Thonstad, T. A. Education and Manpower: Theoretical Models and Empirical Applications. London: Oliver and Boyd, 1969.

Tinbergen, J., and H. C. Bos. "A Planning Model for the Educational Requirements of Economic Development." In Econometric Models of Education: Some Applications. Paris: OECD, 1965, 1-53.

Tinbergen, J., and H. C. Bos. "Appraisal of the Model and Results of Its Application." In Econometric Models of Education: Some Applications. Paris: OECD, 1965, 95-99.

Tracz, G. W. An Overview of Optimal Control Theory Applied to Educational Planning. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, February, 1969.

Washington State Higher Education Facilities Commission. Higher Education Enrollment Projections: A Model. Olympia: WSHEFC, 1970.

Wasik, J. L. The Development of a Mathematical Model to Project Enrollment in a Community College System. Occasional Paper No. 11. Raleigh: Center for Occupational Education, North Carolina State University, 1971.

Wasik, J. L. A Review and Critical Analysis of Mathematical Models Used for Estimating Enrollments in Educational Systems. Technical Paper No. 5. Raleigh: Center for Occupational Education, North Carolina State University, 1971

Webster, J. "The Cohort-Survival Ratio Method in the Projection of School Attendance." Journal of Experimental Education, XXXIX (1970), 89-96.

Wederlin, I. "On the Projection of Numbers of Elementary School Entrants, with Examples from Jordan and Syria." Scientia Paedogogica Experimentalis, III (1966), 43-62.

Williams, G. L. "Planning Models for the Calculation of Educational Requirements for Economic Development - Greece." In Econometric Models of Education: Some Applications. Paris: OECD, 1965, 77-93.

Williams, G. L. "Towards a National Educational Planning Model." Socio-Economic Planning Science, II (1969), 251-261.

Wong, Y. "Computer Simulation of Student Mobility Patterns." Journal of the Association for Educational Data Systems, II (1969), 3-19.

Wurtele, Z.S. Mathematical Models for Educational Planning. SP-3015. Santa Monica, California: System Development Corporation, 1967.

Zabrowski, K. "The Dynamod Model of Student and Teacher Population Growth." Socio-Economic Planning Science, II (1969), 455-464.